THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 53, PART 2 1 March 1941 No. 296

THE SENSITIVITY OF THE EYE TO SMALL COLOUR DIFFERENCES

BY W. D. WRIGHT, A.R.C.S., D.Sc.,
Imperial College of Science and Technology

Received 14 *October* 1940

ABSTRACT. The paper records the measurement of small colour steps for various lines across the C.I.E. colour chart. The whole area of the chart has been covered by the author in a series of observations on 35 lines in various directions on the chart, and six of the more important lines have also been tested by three other observers. The results are illustrated graphically by a series of short lines of the appropriate length drawn on the colour diagram and they are also tabulated. The very wide variation in size of step in different parts of the C.I.E. chart is demonstrated. The results have been transformed to the rectangular uniform-chromaticity system of Breckenridge and Schaub, and a much closer approximation to equality in the length of step is apparent, although a considerable discrepancy still exists in the red corner of the chart. The application of the results to the specification of colour tolerances and to the adoption of an agreed uniform chromaticity scale is discussed.

§1. INTRODUCTION

IN recent years there has been a growing interest in the specification of colour tolerances, a problem as important in some ways as that of colour specification itself. In the C.I.E. system of colour measurement, a colour is specified in terms of three co-ordinates, X, Y and Z, and may be represented in a two-dimensional chart by a point plotted in terms of two of these co-ordinates, say X and Y. In the normal way the axes of these co-ordinates are drawn at right angles to one another. For the purpose of specifying tolerances, it would be convenient if a given distance at any point and in any direction in the colour chart always corresponded to an equal sensory colour difference, but in the C.I.E. system this condition does not hold. Several oblique projections of the C.I.E. chart have been put forward (Judd, 1935 a and b; MacAdam, 1937; Breckenridge and Schaub, 1939), and these uniform-chromaticity scales, as they are called, were approximations to this ideal so far as could be judged from the experimental evidence available at the time they were proposed.

It has always been recognized that further data were necessary before the merits of these systems could be adequately assessed, and this need was specifically

acknowledged in a recommendation of the Commission Internationale de l'Eclairage at its meeting in Holland in 1939. It was partly in response to this recommendation that the work described in this report was undertaken.

The data previously available consisted in the main of measurements of the hue steps round the spectrum locus, measurements on the steps between white and the spectrum locus and results along one or two isolated lines in the colour chart, e.g. along the locus of black-body radiators. To cover the colour field more comprehensively, any fresh investigation should therefore include observations on the line going from blue to red through the purples, on lines from the purples to white, and for various lines across the colour field on either side of the white point. Moreover, it would not be very satisfactory to determine further isolated sets of results and fit them piecemeal into existing data; rather, the whole field should be tested afresh by one group of observers under one set of observing conditions. Only in this way, for example, would it be possible to compare the actual size of a hue step with that, say, of a saturation or a purple step.

A scheme of this magnitude involves a large number of observations of a type that have to be recorded in fairly short sessions to avoid undue fatigue. To limit the work to reasonable proportions, it was therefore planned that the author would cover the colour field with as complete a network of lines as possible and then select a number of the more important lines for testing by further observers. As it was impossible to continue the research at the Imperial College during the summer vacation owing to the war, this programme had to be modified to the extent that the additional observers totalled only three and the number of lines on which they made observations amounted only to six. It may, if necessary, be found possible to supplement these results at a later date, but the future appeared too uncertain to justify delaying the publication of the present results.

§ 2. EXPERIMENTAL ARRANGEMENTS

The apparatus employed was the new model of the author's trichromatic colorimeter, a detailed description of which has already been published (Wright, 1939). In this instrument two spectra, W_1 and W_2, are formed, and prisms are introduced by which the light from narrow strips of one spectrum is reflected into one half of a square field of view of 2° angular subtense, while light from strips of the second spectrum illuminates the other half of the field. For the observations described here, each half was illuminated by a mixture of two similar beams, say monochromatic radiations of wave-lengths λ_1 and λ_2. The colour about which the discrimination sensitivity was to be measured was determined by fixing the proportions of λ_1 and λ_2 in one half of the field by a suitable choice of neutral filters placed in the W_1 spectrum at λ_1 and λ_2. The overall intensity of this half of the field could also be varied over a small range by means of a shallow photometer wedge mounted immediately behind the photometer prism. (This wedge, together with a fixed compensating wedge to keep the field uniform, was the only material addition to the apparatus required for undertaking the present

investigation.) In the other half of the field, the intensities of both λ_1 and λ_2 were continuously variable, using the wedges normally employed in colour-matching observations.

A discrimination step was then recorded as follows :—For convenience, the two halves of the field were first equated in colour and intensity. No reading of this setting was taken, as it merely served to set the apparatus in the " ready " position to make the observation. The amount of, say, λ_1 was then altered until a difference in colour between the two halves of the field was detected, the intensity of the other half of the field being simultaneously adjusted to remove any brightness difference that might otherwise have been interpreted as a colour difference. When a step had been observed, the instrument readings of the photometer wedges controlling λ_1 and λ_2 were taken. The intensity of λ_1 would then again be adjusted but in the opposite direction, until a colour difference on the other side of the mean colour was observed. The new reading for λ_1 was noted and the whole procedure repeated until three settings on either side had been made. The calculations by which the actual step size was derived from the average of these readings are given in the next section.

In the case of observations in which one of the components was a purple, the principle of the experiment was the same, although the practical arrangements were slightly different. In the description of the instrument (Wright, 1939), an auxiliary optical system for the measurement of saturation discrimination was described, in which light from a 1000-watt lamp was focused on to a rotating sector coated with magnesium oxide. In the present case a suitable purple filter was inserted, and, as the sector rotated, both halves of the field were illuminated by a mixture of purple and a selected monochromatic radiation, λ, from the colorimeter proper. To observe a colour step, it was only necessary to vary the intensity of λ in one half of the field and to control the overall intensity in the other. However, when the purple component is small, the intensity variation can be reduced if the colour change is produced by varying the amount of the purple in the mixture, instead of that of the monochromatic component. This could not be conveniently provided for, and as an alternative the operator made a change of approximately the right magnitude by introducing a calibrated neutral filter into one half of the purple beam; the observer then completed the setting by a small adjustment to the intensity of λ. The only disadvantage of this arrangement was that it took slightly longer.

In viewing the field, the observer's head was fixed with his eye-pupil central relative to the exit-pupil of the apparatus, by means of the mouthpiece mounting fitted to the colorimeter. Although this device was primarily intended for binocular matching experiments, it is also used with advantage for monocular observations. In view of the phenomena investigated by Stiles and Crawford, a method of fixing the head might well be used more widely in visual experiments, but it is especially useful in the author's colorimeter, because of the small size of exit-pupil normally employed in the instrument.

With the exception of the observations in the immediate neighbourhood of 0·46 μ, the intensity of the field was maintained at an approximately constant level (about 75 photons retinal illumination for the author, 125 photons for the other three observers) over the whole of the colour chart. Intensity did not appear, however, to have an important effect on the size of step, for, under the given conditions of observation, an increase in intensity by a factor of 10 did not lead to a reduction in the size of step by more than 50 per cent. Actually for any one observer, the intensity did not vary from one observation to another by more than a factor of 2. The level of illumination was governed by the amount of light available at the blue end of the spectrum and, while not very high, was quite adequate for a field viewed against a dark surround by a dark-adapted eye. Owing to the limited intensity, no observations at wave-lengths shorter than 0·46 μ were recorded.

It was decided that in making a step, the criterion used should be rather larger than is usually understood by a just noticeable step. In recording the latter, something akin to a sporting instinct can be introduced in an attempt to discern as minute a difference as possible, but this takes time and imposes a considerable strain on the observer. If the standard is made less exacting, the strain is reduced and a lengthy investigation can be completed more successfully. On the other hand, it is probably less easy to maintain a consistent size for the step ; in the case of the author, for example, the steps recorded in the first two or three weeks of the investigation were appreciably smaller than in the later months. In time, a reasonably settled judgment was attained, and all the observations reported here were recorded after this had been achieved.

It is, of course, quite true that some variations in judgment are inevitable. No doubt the ideal method would be to make a large number of matches at each point in the colour chart and then to analyse the spread of the observations, but in practice this would be an impossibly lengthy process.

§3. CALCULATION OF STEP SIZE

To derive the size of step in terms of a length in the C.I.E. colour chart, one or two preliminary colorimetric observations were necessary. These, and the subsequent calculations, can best be described by means of a numerical example.

For this, we can take some actual observations along the line in the chart joining 0·58 μ and 0·48 μ. The apparatus was first used as a trichromatic colorimeter (Wright, 1939) using the instrument primaries 0·65 μ, 0·53 μ and 0·46 μ, and initial matches were made on 0·5825 μ and 0·4940 μ to establish their units (Wright, 1929). The next observation was a trichromatic colour match of a suitable mixture of 0·58 μ and 0·48 μ, the criterion of suitability being simply that the mixture should be reasonably intermediate between the two limiting wave-lengths and not too near either one or the other. From the match, the trichromatic coefficients of the colour in terms of the three instrument primaries were calculated and thence transformed into C.I.E. units. An actual observation gave

$$C = 0{\cdot}294\,X + 0{\cdot}300\,Y + 0{\cdot}406\,Z.$$

The position of C on the C.I.E. chart was plotted and the point nearest to it on the line joining 0·58 μ and 0·48 μ marked off. The lengths into which the line was divided by this point were measured graphically and gave the distance from 0·48 μ as $l = 0{\cdot}264$ and the distance from 0·58 μ as $l = 0{\cdot}288$, the overall length, as found graphically, being 0·552. The units in which these lengths were measured were the same as those of X and Y; thus if the chart were drawn so that 1 inch corresponded to 0·1 of X or Y, then 1 inch along the (0·48 to 0·58 μ)-line would also be 0·1 units.

The next stage in the experiment was to remove one of the colorimeter primaries and change the wave-lengths of the other two to 0·48 μ and 0·58 μ, an operation that could be readily carried out on the apparatus. The observer then made a new match on C, using the two new primaries alone, the instrument readings were recorded and the corresponding values on the wedge calibration curves noted. From this result a correction factor to be applied to the 0·58 μ values was derived, the factor being chosen to bring the amount of 0·48 μ to 0·58 μ to the ratio of 0·288 to 0·264.

The observer then made his setting for a colour step, first on one side of C and then on the other. The mixture of 0·48 μ and 0·58 μ was then altered and the steps measured about a number of points between the two wave-lengths. In each observation, the wedge readings were noted and the correction factor to the 0·58 μ values applied throughout. For a particular colour in which C_1 was a setting on the blue side and C_2 that on the yellow side, the corrected results (mean of three observations) were

At C_1, amount of 0·48 μ = 794,
amount of 0·58 μ = 74.
At C_2, amount of 0·48 μ = 794,
amount of 0·58 μ = 208.

These numbers are still quite arbitrary in magnitude, but since the appropriate correction has been applied, they show that the ratio of the distance of C_1 from 0·48 μ to that from 0·58 μ was $\frac{74}{794}$ and for C_2 was $\frac{208}{794}$. To find the actual distance of C_1 and C_2 from 0·48 μ, the overall length of 0·552 was multiplied by $\frac{74}{74+794}$ and $\frac{208}{208+794}$ respectively and gave

Distance of C_1 from 0·48 μ = 0·047.
Distance of C_2 from 0·48 μ = 0·115.

The distance from 0·48 μ of the mid-point between C_1 and C_2 was thus

$$l = \tfrac{1}{2}(0{\cdot}047 + 0{\cdot}115) = 0{\cdot}081$$

and the distance from C_1 to C_2 was

$$(0{\cdot}115 - 0{\cdot}047) = 0{\cdot}068.$$

Since C_1 and C_2 were each separated from the mid-point by a colour step, the

distance between C_1 and C_2 corresponded to two steps. The length of a single step, Δl, was therefore given by

$$\Delta l = \frac{0{\cdot}068}{2} = 0{\cdot}034.$$

The values of Δl were then plotted against the corresponding values of l for all the points observed. This gave the curve in figure 1. (As explained later, the author's values of Δl have all been reduced by a factor of 0·47 to bring them into line with those of the other observers. This reduction has therefore been applied to figure 1, so that, for example, the value of Δl at $l = 0{\cdot}081$ is now 0·016 instead of 0·034). A smooth curve was drawn through the points and from this the length of step for any desired value of l could be read off. The step could also

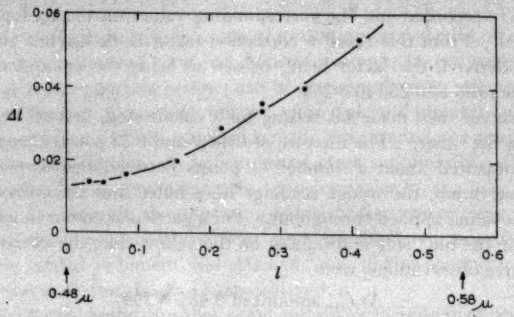


Figure 1. Relation between a colour-discrimination step measured by its length, Δl, in the C.I.E. colour chart and its distance, l, from 0·48 μ in that chart, the steps being observed along the line joining 0·48 μ and 0·58 μ.

be represented on the C.I.E. chart itself by a line of the appropriate length drawn along the (0·48 μ to 0·58 μ)-line and centred about a point at the proper distance from 0·48 μ.

The calculation of the steps along lines spreading from the purple points is no different in principle from the above and need not be described separately. The only additional observation required was the determination of the C.I.E. co-ordinates of the purple colours themselves by an initial trichromatic colour match.

§ 4. RESULTS

In all, 35 lines across the colour chart were tested by the author, including the hue-discrimination curve. The number of points recorded for each line was of the same order as that shown in figure 1, and as a rule the points lay as closely on a smooth curve as in the diagram, although there were exceptions. For each line, a repeat set of observations were made, although owing to the shortage of time only three or four points could usually be repeated. These served, however,

to confirm the shape of the curve and the order of magnitude of the steps. The accuracy of repetition was frequently very good, the results differing from the first set by only a few per cent. In several cases, the difference was greater, of the order of 25 per cent, while for a few of the lines recorded first, the repeat observations, usually obtained some two or three months after the initial set, gave steps

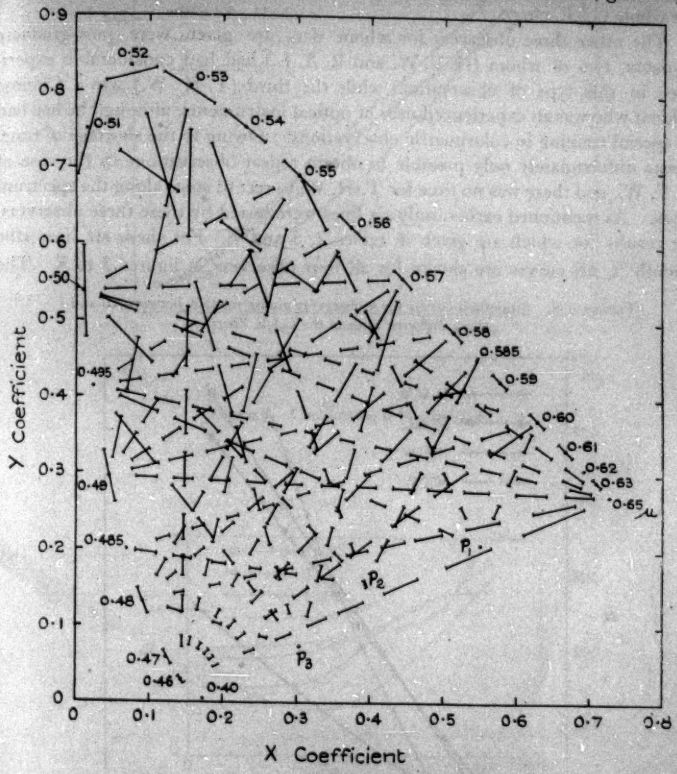


Figure 2. Colour steps shown in the C.I.E. chart by their actual length, Δl, at various points along 35 lines tested by the author.

nearly double the original ones. In these cases, the whole curve was re-observed, as a rule some three or four weeks later, and this third set of observations was generally in good agreement with the second. In only two or three instances was the actual shape of the $\{l, \Delta l\}$ curve not repeated; in the other cases, the changes were nearly always variations of magnitude of the ordinates as a whole. In all but one case, sufficient observations were recorded to establish the true

shape of the curve; for the remaining line, there was no time to trace the nature of the error, and the results have been omitted from the 35 lines reported here.

The author's observations are given in table 1, in which the values of Δl are recorded as derived from the smoothed $\{l, \Delta l\}$ curves. They are also shown graphically in figure 2, in which the size of step has been drawn at intervals across the whole C.I.E. chart.

The other three observers for whom data are given were post-graduate students, two of whom (H. V. W. and R. A. L.) had had considerable experience in this type of observation, while the third (T.-H. W.) was a Chinese student who was an experienced user of optical instruments, although he had had no special training in colorimetric observations. Owing to the shortage of time, it was unfortunately only possible to obtain repeat observations in the case of H. V. W., and there was no time for T.-H. W. to record steps along the spectrum locus. As mentioned earlier, only six lines were tested by these three observers, the results for which are given in tables 2, 3 and 4. For these six lines, the smooth $\{l, \Delta l\}$ curves are shown for all four observers in figures 3 to 8. The

Figures 3–8. Smoothed curves for 4 observers giving relation between Δl and l along different lines in the C.I.E. chart.

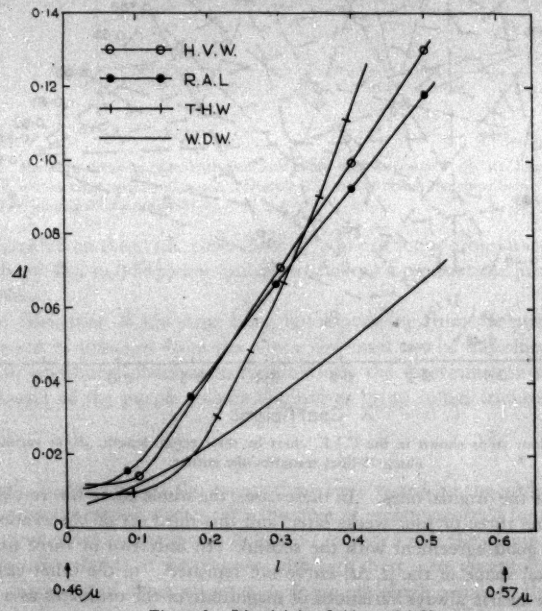


Figure 3. Line joining 0·46 μ and 0·57 μ.

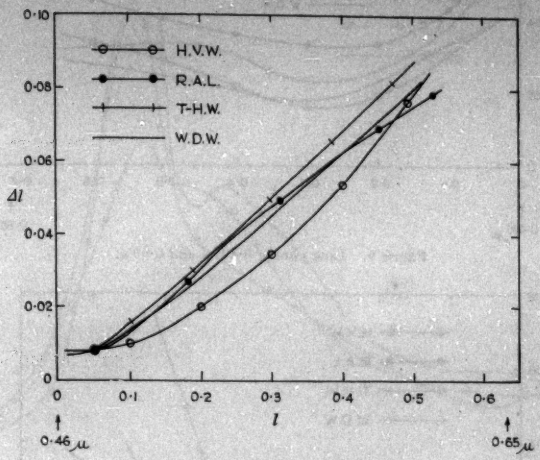


Figure 4. Line joining 0·46 μ and 0·65 μ.

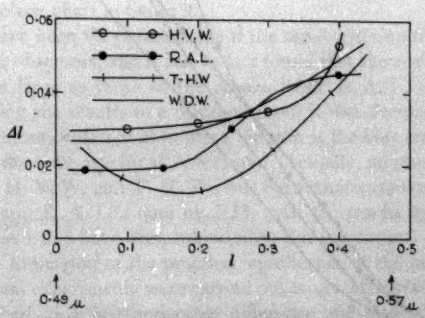


Figure 5. Line joining 0·49 μ and 0·57 μ.

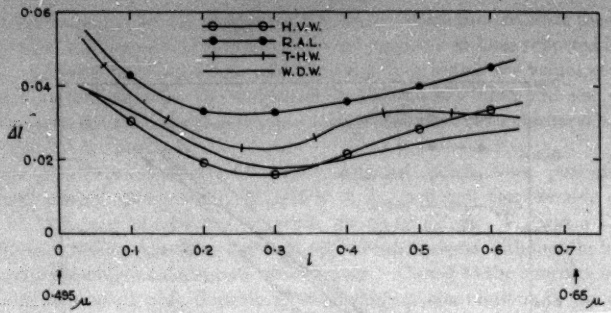


Figure 6. Line joining 0·495 μ and 0·65 μ.

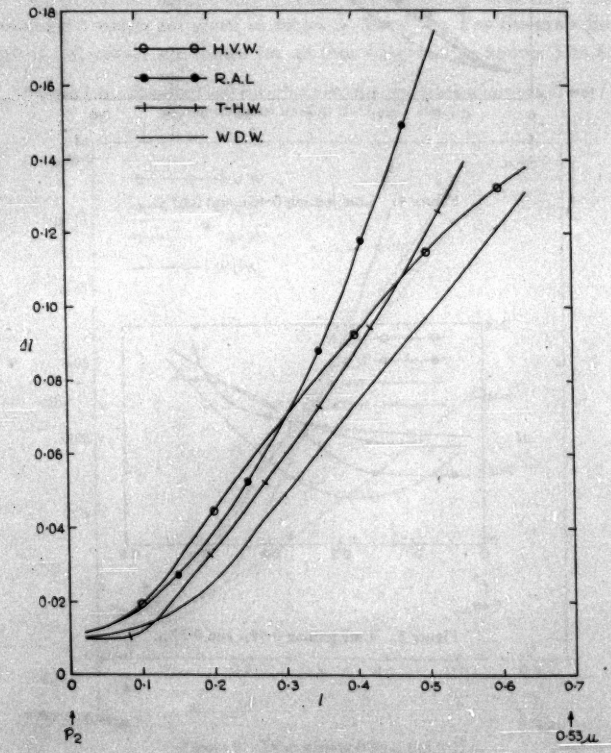


Figure 7. Line joining P_2 and 0·53 μ.

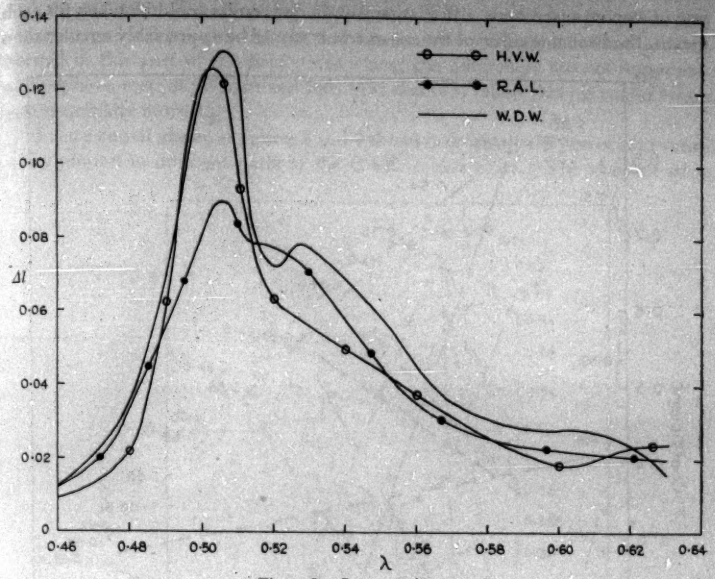


Figure 8. Spectrum locus.

mean data for these lines have been tabulated in table 5, and are shown graphically in the C.I.E. colour chart in figure 9.

It would have been very remarkable if the same criterion for a step had been used by all four observers, and it was in fact found that the author's criterion was about twice as large as those of the others, while R. A. L. made the smallest steps. To bring the results to a common level, it was decided to adjust all the Δl values by factors chosen so as to bring the step at the blue end of the {0·46 μ to 0·65 μ}-line to the same size for all observers. Actually, no change was necessary to the data for H. V. W. and T.-H. W., but the author's results have been multiplied by 0·47 and R. A. L.'s data by 1·33. All the results recorded in figures 2 to 9 and tables 1 to 5 have been subjected to this correction.

As a rough indication of the practical significance of the mean step-size, one or two additional experiments were carried out to correlate this step with what is usually described as a just-noticeable difference and also with the spread of 20 settings aimed at matching a given colour. These showed that the mean step was larger than the just-noticeable difference by a factor of the order of 3 and larger than the overall spread of the matches by a factor of the order of 10. This should mean that when the coefficients of a colour are determined by only 3 or 4 matches, the answer should still be correct to within about $\pm 0{\cdot}05$ of the

size of step reported here. If a larger number of results could be taken for each match, the probable error of the mean result would be appreciably smaller than

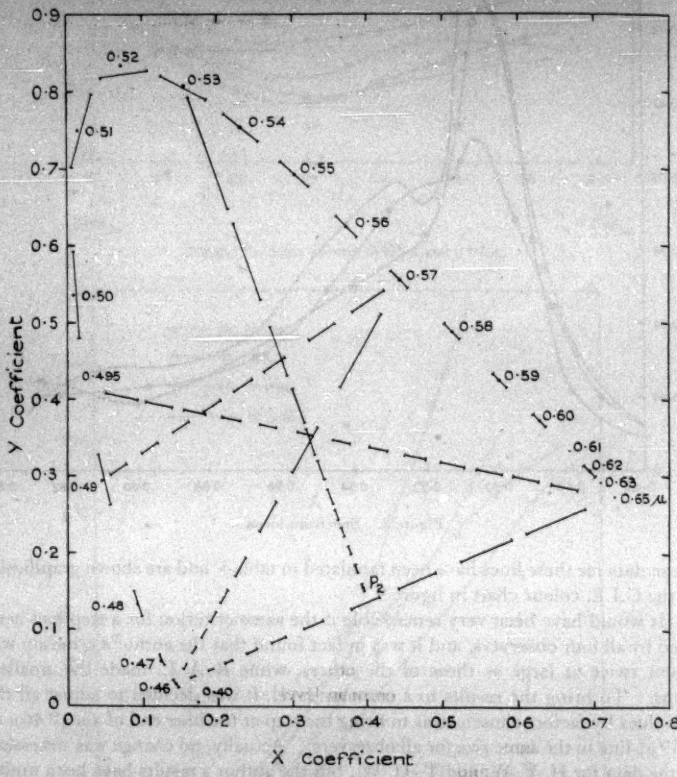


Figure 9. Colour steps for the mean of 4 observers plotted in the C.I.E. chart.

this. This is only true, of course, if instrumental errors can be neglected, and only applies to a colorimeter using a 2° field with a dark surround.

§ 5. DISCUSSION OF RESULTS

In view of the fact that it was possible for the author alone to make a complete series of observations, the data for the other three observers are valuable mainly as an indication of the extent to which the author's colour discrimination can be taken as representative of an ordinary observer. Figures 3 to 8 show that the author's curves do not differ significantly in shape or magnitude from the others,

with the exception of the {0·46 to 0·57 μ}-line. In this, the steps at the yellow end are small, and it is possible that the author's sensitivity differs somewhat from the normal in this part of the field. Elsewhere, the differences are not important, and previous records (Wright and Pitt, 1935 and 1937) show that his colour vision is substantially normal.

Even a casual glance at figures 2 and 9 shows how widely the size of step varies when plotted in different parts of the C.I.E. colour chart. The smallest steps

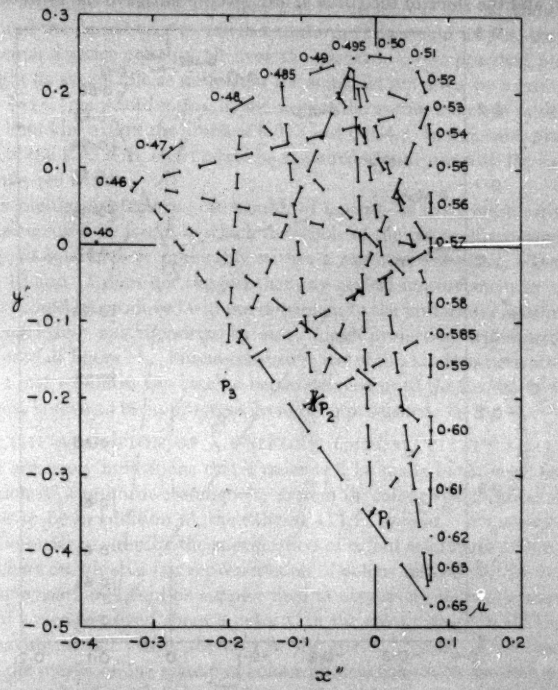


Figure 10. Colour steps plotted in the R.U.C.S. chart for 18 of the lines tested by the author.

are in the blue corner and near the purple line, the longest steps in the green and yellow, and the ratio of the shortest to the longest step is at least 1 : 20.

It has been recognized for some years that there were large variations in sensitivity across the C.I.E. chart, and the development of the so-called uniform-chromaticity scales by Judd (1935 a and b), MacAdam (1937) and Breckenridge and Schaub (1939) was undertaken in an attempt to produce a chart in which the steps were all of the same size. It was of obvious interest, therefore, to transform

the present data into at least one of these scales, and the transformation has been carried out for the rectangular uniform-chromaticity scale (R.U.C.S.) of Breckenridge and Schaub. The transformations were made graphically with the aid of a large-scale diagram of the C.I.E. lattice distorted to the R.U.C.S. framework, a diagram of this type having been kindly given to the author by Mr. J. G. Holmes. For each line, the distances of corresponding points from one end were measured on both the rectangular and the distorted C.I.E. lattice and plotted against one another, and the slope of the curve at any given point gave the ratio of the Δl on

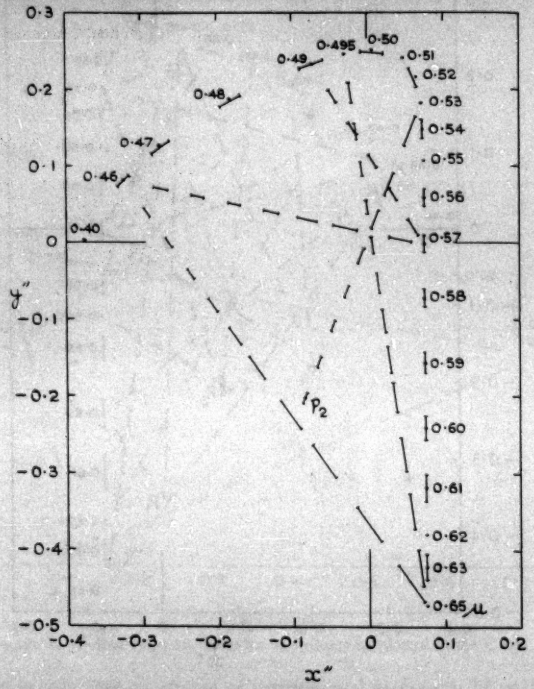


Figure 11. Colour steps for the mean of 4 observers plotted in the R.U.C.S. chart.

the R.U.C.S. system to that on the C.I.E. system. In this way the transformation of 18 of the lines in table 1 and all the lines in table 5 was carried out, the transformed data being given in tables 6 and 7. The steps are also shown graphically on the R.U.C.S. chart in figures 10 and 11.

The very great improvement effected by this transformation is obvious. It is most marked in the manner in which the green, and to a less extent the yellow, lines have been shortened and at the same time the short lines at the blue corner

have been lengthened. On the other hand, in the red corner the lines have been drawn out in one direction and contracted in the transverse direction, and taking the colour chart as a whole, the ratio of the shortest step to the longest is still about 1 : 4. The steps round the spectrum locus also tend to be rather longer than those going inwards from the locus.

It is, of course, quite certain that no linear projection of the C.I.E. diagram can give an exact equality of the colour steps, but it is also true that, considering the limited accuracy to which the steps can be measured, the variation among observers and the varied uses to which a uniform scale would be put, there is no need to demand exact equality all over the diagram. For practical purposes, a scale might be acceptable as uniform if the steps did not differ by a ratio greater than, say, 2 : 1; this would mean, in the case of the author's data in table 6, that the steps should lie within the limits of 0·017 and 0·034. In that case, practically the whole of the R.U.C.S. chart could be regarded as uniform with the exception of the purple-red corner.

This is highly satisfactory. It would, of course, be even more satisfactory if a projection could be found in which the whole of the colour chart came within the suggested tolerance or preferably within a narrower tolerance. Inspection of figures 10 and 11 does not suggest that any radical improvement in one part of the chart could be produced without detriment to the uniformity in other areas, and this conclusion was supported by some rough projection experiments using a transparency of figure 11. But no exhaustive test of this kind has been attempted, and such a test would in any case be better delayed until the results of other investigations, known to be in progress, have been published.

§6. THE ADOPTION OF A UNIFORM CHROMATICITY SCALE

There are many indications that a move will be made in the near future for the adoption of a uniform-chromaticity system of colour specification either to replace, or to be in addition to, the existing C.I.E. system. A uniform scale is desirable in order to simplify the specification of colour tolerances and to provide a colour chart on which a fair representation of colour differences can be shown. For example, nothing could be simpler than to express a colour tolerance by the radius r of a circle centred about a point C in the colour diagram, all acceptable colours having to fall within the area of the circle. Again, in comparing, for example, the merits of one system of colour photography with another, it may be convenient to plot the colour errors in the colour chart. Inevitably, the magnitude of the error is associated with the spread of the points on the chart, yet this may produce a gross misconception unless something approximating to a uniform-chromaticity chart is employed.

It is true that even a uniform-chromaticity scale will not always give a true picture of the subjective importance of colour errors ; thus in some cases hue errors may be more serious than errors of saturation and vice versa. Nor can it take into account the vital part that manufacturing considerations must play in the specification of colour tolerances. Nevertheless, in many problems the

uniform scale could be applied directly and it seems reasonably certain that in all those cases in which special considerations arose, the significance of the tolerances would be more accurately appreciated with a uniform scale as the basic system than with a non-uniform scale.

MacAdam (1939) has suggested that even if the principle of a uniform-chromaticity scale were adopted, several different charts would be necessary to meet the needs of various industries and conditions of observation. If this were the case, the whole position would be much less favourable, as it is doubtful whether a system comprising several alternative ways of colour specification could ever be operated successfully in industry. A more acceptable method would probably be to have one uniform scale and allow for the different needs by the use of different tolerance areas, e.g. circles, ellipses, triangles, rectangles etc. It may also be that the effect of the conditions of observation on the distribution of the steps across the colour chart has been exaggerated. Undoubtedly the magnitude is affected, for example, by size of field, but there is little evidence to suggest that the relative size of step in different parts of the chart would be altered to an extent sufficient to call for a change in the uniform-chromaticity scale. Further experiment is admittedly called for before this conclusion can be verified.

Any change from the existing C.I.E. system is to be deplored unless something really concrete is gained in the process. The use of a distorted C.I.E. lattice for graphical colour records has been suggested by Holmes (1940), and this would in some measure avoid the necessity of substituting a new system for the C.I.E. system. Yet if an effectively uniform scale could be found, then in fairness to those who will be working on colorimetry during the next decades, a change to an entirely new co-ordinate system seems to be the logical step to take.

The measurement of just-noticeable differences across the colour chart is a necessary preliminary to the derivation of a satisfactory uniform-chromaticity scale, but it is a mistake to think that it provides all the information required. For example, even if we are prepared to accept the assumption that just-noticeable differences correspond to equal sense differences or intervals [not an *a priori* assumption, see, for example, Boring (1939), Wright (1938)], it is certainly not legitimate to integrate these differences and conclude that distances of some length across the chart will correspond to equal sense differences. The spacing of large differences would call for a special and very difficult investigation of its own. Again, the present results have been recorded at a practically uniform intensity level, and although this is the most straightforward way to control the intensity, it may not approximate very closely to practical requirements. Thus in the case of pigments or dyes, saturated blues, blue-greens or purples will generally be very considerably darker than reds, yellows or greens of similar concentrations. We are naturally led to the conclusion that a comprehensive investigation would be three-dimensional, involving intensity differences as well as colour differences, a gigantic task that is to some extent in progress in connection with the Munsell system of colour specification (Newhall, 1939).

In tables 1–7, the size of the colour step is given along various lines in the C.I.E. and R.U.C.S. colour charts. The position at which the step has been measured is given by l, the distance in the appropriate chart from the shorter of the two wave-lengths at the extremity of the line. In the case of a line through one of the purple points, l is measured from that point.

The C.I.E. co-ordinates of the three purples used are

$$P_1 = 0{\cdot}550X + 0{\cdot}206Y + 0{\cdot}244Z,$$
$$P_2 = 0{\cdot}395X + 0{\cdot}155Y + 0{\cdot}450Z,$$
$$P_3 = 0{\cdot}304X + 0{\cdot}074Y + 0{\cdot}622Z.$$

Table 1. Steps in the C.I.E. chart. Observations by author

l		0·05	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80
Line 0·46 −0·53 μ		0·015	0·020	0·035	0·049	0·066	0·089	0·118		
0·46 −0·55		0·010	0·015	0·030	0·053	0·077	0·103	0·127		
0·46 −0·57		0·008	0·012	0·023	0·036	0·052	0·068			
0·46 −0·585		0·009	0·014	0·026	0·038	0·052	0·070			
0·46 −0·60		0·008	0·013	0·024	0·036	0·048	0·062			
0·46 −0·65		0·008	0·014	0·028	0·045	0·063	0·080			
0·48 −0·56		0·025	0·026	0·035	0·046	0·055	0·061			
0·48 −0·58		0·014	0·016	0·026	0·038	0·054	0·073			
0·48 −0·65		0·012	0·014	0·019	0·028	0·037	0·050	0·063		
0·485−0·60		0·018	0·016	0·015	0·019	0·027	0·036			
0·49 −0·53		0·054	0·062	0·080	0·100	0·126				
0·49 −0·55		0·031	0·034	0·042	0·051	0·066				
0·49 −0·57		0·027	0·027	0·030	0·038	0·049				
0·49 −0·59		0·037	0·032	0·022	0·023	0·034	0·045			
0·49 −0·65		0·028	0·026	0·024	0·024	0·029	0·035	0·044		
0·495−0·56		0·023	0·023	0·023	0·029	0·035				
0·495−0·58		0·027	0·025	0·023	0·023	0·023				
0·495−0·65		0·038	0·033	0·023	0·017	0·020	0·024	0·027		
0·50 −0·55		0·031	0·031	0·031	0·031					
0·50 −0·57		0·043	0·039	0·034	0·031	0·029				
0·50 −0·60		0·053	0·046	0·034	0·023	0·017	0·016	0·018		
0·50 −0·65		0·052	0·045	0·033	0·024	0·022	0·022	0·023	0·026	
0·51 −0·65		0·104	0·094	0·075	0·058	0·042	0·028	0·023	0·024	0·030
P_1−0·495		0·019	0·020	0·024	0·031	0·041	0·052			
P_1−0·55		0·014	0·016	0·023	0·034	0·052	0·072			
P_1−0·60		0·018	0·022	(l=	0·15,	Δl=	0·026)			
P_2−0·485		0·016	0·014	0·019	0·029					
P_2−0·50		0·015	0·015	0·022	0·038	0·067	0·099			
P_2−0·53		0·011	0·014	0·025	0·049	0·074	0·097	0·122		
P_2−0·585		0·031	0·028	0·038	0·062					
P_3−0·48		0·010	0·010	(l=	0·15,	Δl=	0·014)			
P_3−0·495		0·007	0·014	0·032	0·062					
P_3−0·52		0·012	0·017	0·031	0·052	0·078	0·105	0·134	0·170	
P_3−0·57		0·018	0·021	0·031	0·053	0·085				
Spectrum locus	λ	0·46μ	0·47	0·48	0·49	0·50	0·51	0·52	0·53	0·54
	Δl	0·013	0·021	0·036	0·069	0·120	0·120	0·072	0·077	0·068
	λ	0·55	0·56	0·57	0·58	0·59	0·60	0·61	0·62	0·63
	Δl	0·058	0·045	0·035	0·031	0·028	0·028	0·027	0·024	0·016

Table 2. Steps in the C.I.E. chart. Observations by H. V. W.

l		0·05	0·10	0·20	0·30	0·40	0·50	0·60		
Line 0·46 —0·57 μ		0·012	0·016	0·042	0·071	0·100	0·130			
0·46 —0·65		0·008	0·010	0·020	0·035	0·054	0·078			
0·49 —0·57		0·030	0·030	0·032	0·035	0·054				
0·495—0·65		0·048	0·038	0·026	0·023	0·030	0·033	0·032		
P_2—0·53		0·013	0·020	0·045	0·070	0·093	0·115	0·132		
Spectrum locus	λ	0·46 μ	0·47	0·48	0·49	0·50	0·51	0·52	0·53	0·54
	Δl	0·009	0·013	0·022	0·060	0·123	0·098	0·063	0·056	0·050
	λ	0·55	0·56	0·57	0·58	0·59	0·60	0·61	0·62	0·63
	Δl	0·044	0·037	0·031	0·026	0·021	0·018	0·020	0·023	0·024

Table 3. Steps in the C.I.E. chart. Observations by R. A. L.

l		0·05	0·10	0·20	0·30	0·40	0·50	0·60		
Line 0·46 —0·57 μ		0·012	0·018	0·043	0·068	0·093	0·118			
0·46 —0·65		0·008	0·013	0·031	0·048	0·062	0·075			
0·49 —0·57		0·019	0·019	0·024	0·039	0·045				
0·495—0·65		0·052	0·043	0·033	0·033	0·035	0·040	0·045		
P_2—0·53		0·013	0·019	0·039	0·070	0·112	0·166			
Spectrum locus	λ	0·46μ	0·47	0·48	0·49	0·50	0·51	0·52	0·53	0·54
	Δl	0·012	0·019	0·034	0·056	0·081	0·083	0·077	0·070	0·057
	λ	0·55	0·56	0·57	0·58	0·59	0·60	0·61	0·62	0·63
	Δl	0·045	0·035	0·029	0·026	0·024	0·022	0·022	0·021	0·020

Table 4. Steps in the C.I.E. chart. Observations by T.-H. W.

l	0·05	0·10	0·20	0·30	0·40	0·50	0·60
Line 0·46 —0·57 μ	0·009	0·010	0·027	0·065	0·115		
0·46 —0·65	0·008	0·015	0·033	0·050	0·069	0·087	
0·49 —0·57	0·022	0·016	0·013	0·023	0·043		
0·495—0·65	0·037	0·030	0·019	0·016	0·021	0·028	0·033
P_2—0·53	0·010	0·012	0·034	0·060	0·088	0·120	

Table 5. Steps in the C.I.E. chart. Mean of 4 observers

l		0·05	0·10	0·20	0·30	0·40	0·50	0·60		
Line 0·46 —0·57 μ		0·010	0·014	0·034	0·060	0·091	0·122			
0·46 —0·65		0·008	0·013	0·028	0·045	0·062	0·080			
0·49 —0·57		0·024	0·023	0·025	0·034	0·048				
0·495—0·65		0·044	0·036	0·025	0·022	0·026	0·031	0·034		
P_2—0·53		0·012	0·016	0·036	0·062	0·092	0·124	0·160		
Spectrum locus	λ	0·46 μ	0·47	0·48	0·49	0·50	0·51	0·52	0·53	0·54
	Δl	0·011	0·018	0·031	0·062	0·108	0·100	0·071	0·068	0·065
	λ	0·55	0·56	0·57	0·58	0·59	0·60	0·61	0·62	0·63
	Δl	0·049	0·039	0·032	0·028	0·024	0·022	0·023	0·023	0·020

Table 6. Steps in the R.U.C.S. chart. Observations by author

l		0·05	0·10	0·20	0·30	0·40	0·50	0·60	0·70	
Line 0·46 —0·53μ		0·022	0·022	0·022	0·023	0·020				
0·46 —0·57		0·017	0·016	0·016	0·019	0·022				
0·46 —0·60		0·016	0·016	0·018	0·027	0·030	0·030			
0·46 —0·65		0·014	0·014	0·021	0·035	0·048	0·055	0·057		
0·48 —0·58		0·015	0·015	0·020	0·025					
0·49 —0·53		0·025	0·024							
0·49 —0·57		0·021	0·019	0·019						
0·495—0·58		0·017	0·016	0·016	0·016					
0·495—0·65		0·024	0·020	0·017	0·021	0·027	0·032	0·037	0·042	
P_1—0·495		0·033	0·030	0·027	0·027	0·030	0·034			
P_1—0·60		0·015	0·015							
P_2—0·485		0·024	0·021	0·018	0·022	0·026				
P_3—0·53		0·020	0·016	0·017	0·026	0·028				
P_4—0·585		0·024	0·018	(l=	0·15,	Δl=	0·021)			
P_5—0·48		0·025	0·022	0·019	0·017					
P_6—0·495		0·018	0·018	0·019	0·029	0·042				
P_7—0·57		0·021	0·021	0·021	0·025					
Spectrum locus	λ	0·46μ	0·47	0·48	0·49	0·50	0·51	0·52	0·53	0·54
	Δl	0·024	0·033	0·036	0·035	0·030	0·027	0·029	0·030	0·027
	λ	0·55	0·56	0·57	0·58	0·59	0·60	0·61	0·62	0·63
	Δl	0·026	0·027	0·025	0·026	0·031	0·037	0·042	0·040	0·032

Table 7. Steps in the R.U.C.S. chart. Mean of 4 observers

l		0·05	0·10	0·20	0·30	0·40	0·50	0·60	0·70	
Line 0·46 —0·57 μ		0·019	0·017	0·020	0·032	0·027				
0·46 —0·65		0·014	0·014	0·021	0·035	0·048	0·055	0·057		
0·49 —0·57		0·019	0·018	0·018	(l=	0·25,	Δl=0	·024)		
0·495—0·65		0·027	0·021	0·019	0·026	0·035	0·042	0·047	0·049	
P —0·53		0·014	0·015	0·021	0·031	0·043				
Spectrum locus	λ	0·46μ	0·47	0·48	0·49	0·50	0·51	0·52	0·53	0·54
	Δl	0·019	0·028	0·032	0·033	0·031	0·024	0·027	0·028	0·024
	λ	0·55	0·56	0·57	0·58	0·59	0·60	0·61	0·62	0·63
	Δl	0·022	0·022	0·022	0·023	0·027	0·032	0·036	0·038	0·033

§7. ACKNOWLEDGEMENTS

I should like once again to express my gratitude to the Medical Research Council for their continued financial support. I have also to thank the three observers who provided very useful data under rather trying conditions, more especially Mr. H. V. Walters, who recorded my own observations with care and patience and helped generally in the course of the experiments.

REFERENCES

BORING, E. G., 1939. *Amer. J. Psychol.* **52**, 384.
BRECKENRIDGE, F. C. and SCHAUB, W. R., 1939. *J. Opt. Soc. Amer.* **29**, 370.
HOLMES, J. G., 1940. *Proc. Phys. Soc.* **52**, 359.
JUDD, D. B., 1935 a. *Bur. Stand. J. Res., Wash.*, **14**, 41.
JUDD, D. B., 1935 b. *J. Opt. Soc. Amer.* **25**, 24.
MACADAM, D. L., 1937. *J. Opt. Soc. Amer.* **27**, 294.
MACADAM, D. L., 1939. *Amer. J. Psychol.* **52**, 412.
NEWHALL, S. M., 1939. *Amer. J. Psychol.* **52**, 394.
SMITH, T. and GUILD, J., 1931–2. *Trans. Opt. Soc.* **33**, 73.
WRIGHT, W. D., 1929. *Trans. Opt. Soc.* **30**, 141.
WRIGHT, W. D., 1938. *The Perception of Light* (London : Blackie).
WRIGHT, W. D., 1939. *J. Sci. Instrum.* **14**, 10.
WRIGHT, W. D. and PITT, F. H. G., 1935. *Proc. Phys. Soc.* **47**, 205.
WRIGHT, W. D. and PITT, F. H. G., 1937. *Proc. Phys. Soc.* **49**, 329

A NOTE ON CONCENTRATED FERROMAGNETIC AMALGAMS

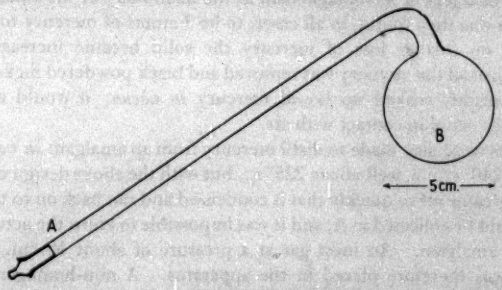
By L. F. BATES, D.Sc.,
University College, Nottingham

Received 4 *November* 1940

ABSTRACT. Attempts to obtain concentrated fluid amalgams of nickel and mercury are described. Concentrated amalgams formed by low-temperature distillation become ferromagnetic when their constitution reaches 3 atoms of mercury to 1 of nickel.

§ 1. DESCRIPTION OF EXPERIMENTS

As it was recently shown (Bates and Baker, 1940) that dilute amalgams of nickel in mercury exhibit ferromagnetism when once they have been heated to a temperature above 225° c., it was thought worth while to find if such amalgams could be concentrated to make, as it were, a kind of ferromagnetic liquid. It is not convenient to prepare by electrolysis amalgams



Distillation apparatus.

containing much more than 2 per cent. of nickel by weight, as they tend to become increasingly solid with rise in nickel content, and consequently, amalgams so prepared were concentrated by removal of mercury by distillation *in vacuo*.

The freshly prepared amalgam was carefully washed in distilled water, roughly dried and placed in the bulb B of an apparatus, of the form shown in the figure, which was evacuated by a Hyvac pump, B being placed in boiling water during the final stages of the evacuation. In this way the amalgam was denuded of any remaining traces of water, and the apparatus was sealed by the simple

device described by Bates, Baker and Meakin (1940). B was then placed in a vessel containing oil which was heated on a sand bath, while the remote end of A was placed in ice. The oil was kept at a chosen temperature for long intervals of time, and the mercury driven off from the amalgam was collected in A. The latter was of fairly uniform bore, so that the volume of mercury driven off could be periodically measured by the length of the column in A, and, accordingly, the composition of the amalgam left in B could be determined from time to time.

The chosen temperatures in these experiments, in each of which a fresh amalgam was used, were 170, 195, 215 and 223° c., all below the temperature 225° c. at which the liquid amalgams become ferromagnetic. At each of these temperatures, and while it still continued to lose mercury, the amalgam in B became a hard matt grey solid at the same composition, viz. 10 atoms of mercury to 1 of nickel. The strange feature about this solid was that it became plastic, one might say fluid, at lower temperatures. It was possible to heat an amalgam of this composition to a temperature at least as high as that at which it was formed and then cool it many times without change in its properties becoming apparent. When maintained for long periods at any one of the above temperatures, the solid amalgam lost more mercury and became increasingly hard and difficult to break up by shaking against the walls of the bulb, and, eventually, a black ferromagnetic powder formed on the surface. The powder was easily shaken from the surface of the amalgam, and its formation coincided with the appearance of traces of ferromagnetism in the main body of the amalgam, whose constitution was then found, in all cases, to be 3 atoms of mercury to 1 of nickel. Thereafter, on further loss of mercury the solid became increasingly ferromagnetic, until all the mercury was removed and black powdered nickel remained. The latter readily soaked up liquid mercury *in vacuo*; it would not soak up mercury after being in contact with air.

Attempts were also made to distil mercury from an amalgam *in vacuo* when B was kept at 240° c., i.e. well above 225° c., but with the above design of apparatus the mercury came off so quickly that it condensed and ran back on to the amalgam before it could be collected in A, and it was impossible to know the actual temperature of the amalgam. An inert gas at a pressure of about 10 cm. of mercury at 240° c. was therefore placed in the apparatus. A non-homogeneous ferromagnetic fluid amalgam was then formed. On evacuation at 240° c. mercury distilled over, leaving first a matt grey ferromagnetic solid and later black powdered nickel.

Nickel powder prepared in these experiments was found to have a specific magnetization of approximately 36 c.g.s. units per gm. in a field of 10,000 oersteds, and a remanent magnetization of approximately 11 c.g.s. units per gm. after it had been in contact with air. The corresponding data for pure solid nickel (hard-drawn) are 57·6 and 33 c.g.s. units per gm. respectively. It was not possible to measure the ferromagnetic properties of a fluid or plastic specimen obtained on cooling to room temperature an amalgam of concentration 10 atoms of mercury

to 1 of nickel, without the use of special apparatus, but we can readily estimate the saturation magnetic moment of the material. One gm. of the fluid contains 0·0286 gm. nickel, so that its saturation intensity of magnetization, taken to be equal to that of the black powder, would be about 14 c.g.s. units per c.c. of fluid, compared with 500 for pure solid nickel. Naturally, its initial susceptibility is low.

§2. ACKNOWLEDGEMENTS

It is with much pleasure that I record my thanks to Mr. E. J. Adamson for his valuable help with the experiments.

REFERENCES

BATES L. F. and BAKER C. J. W., 1940. *Proc. Phys. Soc.* **52**, 536.
BATES L. F., BAKER C. J. W. and MEAKIN, R., 1940. *Proc. Phys. Soc.* **52**, 429.

THE BAND SPECTRUM OF GERMANOUS SULPHIDE IN EMISSION

By R. F. BARROW,
Departments of Inorganic Chemistry and Physics,
Imperial College, London, S.W. 7

Communicated by W. Jevons. Received 2 *November* 1940

ABSTRACT. The band spectrum of germanous sulphide, hitherto known only in absorption, has been excited in emission by means of a heavy-current uncondensed discharge through a mixture of sulphur and germanium dioxide contained in a silica tube, in presence of aluminium. Some 20 new bands lying in the region λ3350–3750 have been measured and assigned to the long-λ end of the main ($A\rightleftarrows X$) system of GeS. Analysis of the data on higher vibrational levels of the ground state derived from these bands confirms the values of ω_e'' and $x_e''\omega_e''$ previously determined from measurements of the system in absorption.

EXPERIMENTAL work aimed at elucidating the general trends of the molecular spectroscopic constants of the diatomic oxides, sulphides, selenides and tellurides of carbon, silicon, germanium, tin and lead has reached the point where certain data are available for nineteen of the twenty molecules of this group (Barrow and Jevons, 1940). Some of these data, however, are based upon experimental observations deficient at least in extent, if not in accuracy. It was therefore proposed to attempt the excitation by positive-column discharges of the emission spectra of all those molecules of this group whose spectra had yet been obtained only in absorption, in the hope of increasing the reliability of the ground-state constants of the molecules in question. Owing to the war, it became necessary to discontinue this plan, but not before a preliminary investigation of the emission spectrum of one member, germanous sulphide, had been made. Vibrational analyses of two ultra-violet band systems in absorption by GeS vapour were carried out by Shapiro, Gibbs and Laubengayer (1932): the present note is concerned with additional bands observed in emission which have been assigned to the long-λ end of the main ($A\rightleftarrows X$) system.

The emission spectrum of GeS was excited by means of a heavy-current uncondensed discharge through a mixture of sulphur and germanium dioxide, in presence of aluminium, contained in a silica tube of a type previously described (Barrow, 1941). The electrical conditions and the spectrograph employed were the same as those used in other investigations in this series.

To face page 117

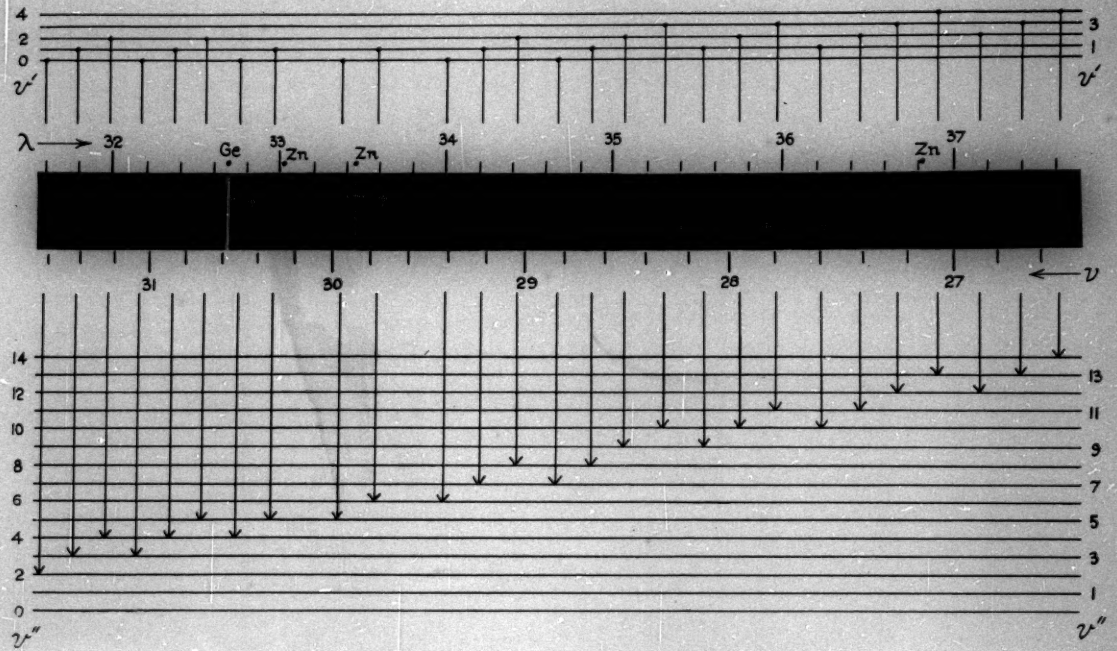


Figure 1. Bands of the $A \rightleftarrows X$ system of GeS in emission (2·4-m. grating, first order).

The plates show a strong development of the $A \rightleftarrows X$ system of GeS in the region λ2850–λ3750, and although at wave-lengths below about λ3100 there is a certain amount of overlapping by bands of the ultra-violet system of SiS, simultaneously excited in the absence of a sufficient excess of aluminium, there are no bands of impurities present in the region λ3350–λ3750 where the new bands are observed. Perhaps because of the development of the SiS system, no trace of the $B \rightleftarrows X$ system of GeS has been found on the plates. Atomic lines present are due to Ge, Si, Al, Ni and Zn. A part of a grating spectrogram showing the main GeS system to the long-λ end of the system-origin is reproduced in figure 1, from which it will be seen that the bands in this region are evenly

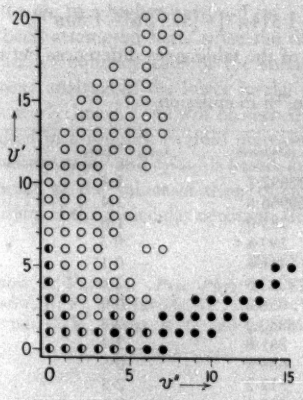


Figure 2.

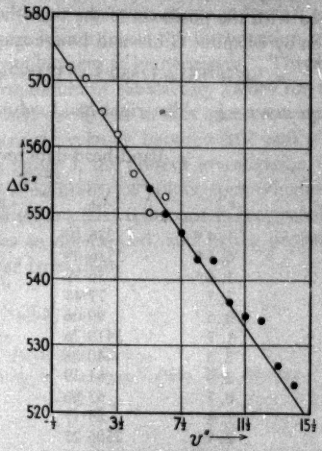


Figure 3.

Figure 2. Distributions of bands in the $A \rightleftarrows X$ system of GeS as observed in emission and in absorption. Black circles represent bands observed in emission only; open circles represent bands observed in absorption only; half-black circles represent bands observed both in emission and in absorption.

Figure 3. The variation of vibrational interval with vibrational quantum number in the ground state of GeS. Black circles represent meaned emission data; open circles represent meaned absorption data from measurements of Shapiro, Gibbs and Laubengayer.

spaced, and form short sequences. Owing to overlapping, many bands are probably lost in the background of partially resolved rotational lines of members of contiguous sequences. The distributions of bands as observed in emission and in absorption are made clear in figure 2: it must only be recalled that measurements on the available plates of the system in emission are probably curtailed on the short-λ side by bands of the SiS system.

The vibrational isotope effect is apparent, even in the reproduction, in that bands increasingly distant from the system origin become increasingly diffuse

in appearance under low dispersion. The effect was examined in detail by the previous investigators of this band system, who were able to recognize heads due to $Ge^{70}S$, $Ge^{72}S$ and $Ge^{76}S$, in addition to those due to the most abundant molecule, $Ge^{74}S$, in most of the bands that they measured. In the majority of the new bands, observed in emission only, four heads due to these isotopic species have also been measured. In a few outlying bands, the calculated isotopic displacements were useful in confirming the v', v'' assignments, but apart from this, the isotope effect here presents no novel or interesting features, and only $Ge^{74}S$ band-head data are included in the table, where the vibrational quantum numbers, wave-lengths and wave-numbers of the new bands are set out. These are compared with the positions of the band-heads as extrapolated from the expression given by Shapiro, Gibbs and Laubengayer, viz.:

$$\nu_{\text{head}}(Ge^{74}S) = 32889{\cdot}5 + [374{\cdot}99u' - 1{\cdot}514u'^2] - [575{\cdot}8u'' - 1{\cdot}80u''^2],$$

where $u = v + \frac{1}{2}$. The right-hand column of the table gives differences between

Band-head data for $Ge^{74}S$ in emission

v', v''	λ_{air}	$\nu_{vac.}$ (observed)	$\nu_{obs.} - \nu_{calc.}$
1, 4*	3236·10	30892·5	− 1·4
0, 5*	3336·09	29966·6	2·4
1, 6*	56·59	783·6	1·6
2, 7	77·43	599·8	− 0·5
0, 6	99·06	411·5	− 0·1
1, 7	3419·76	233·5	2·1
2, 8	40·86	054·2	0·9
3, 9	61·59	28880·2	0·8
0, 7	63·80	861·8	2·4
1, 8	85·17	684·8	0·5
2, 9	3506·27	512·2	2·3
3, 10	28·15	335·4	− 0·6
1, 9	52·57	140·6	− 0·3
2, 10	74·11	27971·0	0·9
3, 11	96·1	799	− 1
1, 10	3622·6	597	− 4
2, 11	44·06	434·1	0·2
3, 12	66·62	265·3	− 1·9
4, 13	89·79	094·1	− 6·9
5, 14	3712·7	26927	− 9
2, 12	16·47	899·6	− 1·7
3, 13	39·84	731·5	− 6·7
4,14	63·0	567	−10
5,15	86·4	403	−11

* The measurements of Shapiro, Gibbs and Laubengayer extend on the long-λ side to the (1, 6) band at λ3356, but both this band and the adjacent (0, 5) are listed as being of very low intensity, and measurements of these bands in emission are therefore included. The (1, 4) band was not observed in absorption.

the observed positions of band-heads and those calculated from the above expression: the agreement is seen to be satisfactory, although it appears that differences larger than those accounted for by purely experimental error are associated with the highest observed vibrational levels in the ground state. This is illustrated graphically in figure 3, where $\Delta G''$ is plotted against v'', meaned emission and absorption data being included. The straight line corresponds to the values of ω_e'' and $x_e''\omega_e''$. Having due regard to the fact that the points to be weighted most heavily are those of lowest v'' values, the line is seen to fit satisfactorily the observed relation between $\Delta G''$ and v'', at least over the range $v''=0$ to $v''=11$. At higher values of v'', the vibrational intervals appear to decrease somewhat less rapidly than the simple equation above indicates, and for the arithmetical representation of these higher levels the introduction of a small term in a higher power of $(v''+\frac{1}{2})$ would seem to be necessary. This has not been attempted here, since the determination of satisfactory values for the vibration frequency and for the anharmonic coefficient in this state rests upon accurate analysis of the lower levels, where the relation between $\Delta G''$ and v'' is very nearly linear. As will be seen from figure 3, the present investigation has resulted in a roughly twofold increase in the number of points upon which the evaluation of ω_e'' and $x_e''\omega_e''$ is based, and there is a related increase in the probable accuracy of the values of these quantities so determined, and in the probable accuracy of the molecular constants derived from them.

REFERENCES

BARROW, R. F., 1941. *Proc. Phys. Soc.* **53**, 40.
BARROW, R. F., and JEVONS, W., 1940. *Proc. Phys. Soc.* **52**, 534.
SHAPIRO, C. V., GIBBS, R. C. and LAUBENGAYER, A. W., 1932. *Phys. Rev.* **40**, 354.

REVERSAL OF RECTIFICATION BY HEAT IN CRYSTAL CONTACTS AND ITS ANALOGY TO A THERMOELECTRIC PHENOMENON

By M. A. EL SHERBINI, M.Sc., Ph.D. and Y. L. YOUSEF, B.Sc.,
Physics Department, Faculty of Science, Fouad I University, Cairo

Received 25 *November* 1940

ABSTRACT. The reversal of rectification previously observed by the authors in molybdenite-copper contacts at low temperatures led them to expect reversal in other rectifying crystals when the temperature is changed over a wide range. Experiments on bornite show that, for voltages less than 1 volt, reversal occurs below 200° c. Silicon is expected to reverse at some temperature below −75° c. for potential differences of the order of several volts. The analogy between the behaviour of rectifying crystals towards heat and the behaviour of a thermocouple whose cold junction is maintained at different temperatures suggests that the rectifying action may be, at least in part, of a thermoelectric origin.

§1. INTRODUCTION

In a previous publication (El Sherbini and Yousef, 1939) the authors found that, when the temperature of a molybdenite-copper contact is raised above −70° c., rectification decreases until it vanishes and reverses its direction between −50 and 0° c. After reversal, rectification increases as the temperature is raised to 100° c.

Crystals which show a marked decrease in rectification by heat behave, in a way, like molybdenite at low temperatures before reversal. The question now arises whether further heating in such specimens causes reversal. Actually, former investigators found temperatures at which the rectifying action of pyrites and galena practically disappears, but these experimenters did not seem to have carried out the heating beyond these temperatures, or to have used voltages favourable for the phenomenon (Flowers 1909; Khastgir and Das Gupta, 1935).

On the other hand, crystals which show a marked increase in rectification with rise of temperature behave similarly to molybdenite after reversal, and are thus likely to reverse when cooled to a sufficiently low temperature. As an example of this class we mention silicon-copper contacts in the interval −75 to 30° c. when the voltage is of the order of several volts. Accordingly, reversal in silicon may be expected below −75° c.

At present we are not in a position to make any low-temperature experiments on silicon to confirm this conclusion, and our work has mainly been concerned with bornite, whose behaviour suggests the probability of reversal at high temperatures.

§ 2. EXPERIMENTAL WORK

The apparatus used is an air-thermostat whose temperature can be electrically controlled between wide limits. The crystal is suspended within the thermostat from a copper holder, which forms one lead for the current. To maintain a constant contact-pressure throughout all the measurements, the whisker is made in the form of a weighted pin of copper which presses freely on the crystal and serves as the other lead for the current. The crystal circuit is of the ordinary type.

In high-temperature work it is necessary to use small voltages in order to avoid the liability of the crystal to breakdown.

§ 3. RESULTS

Within the working limits, bornite changes the sign of rectification. The inversion temperature depends on the applied voltage and on the particular point of contact, but in all the cases studied it is well below 200° c. for voltages less than

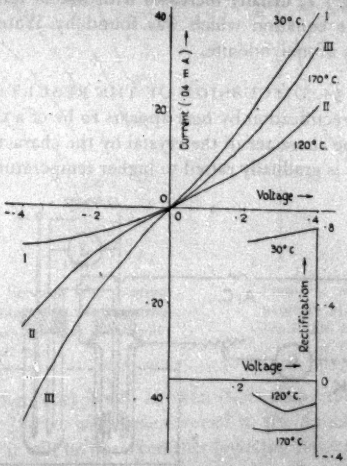


Figure 1. D.C. characteristics and rectification curves of bornite.

1 volt. The employment of higher voltages at high temperatures destroys the rectifying action. Rectification is not restored when the voltage is reduced, even if the temperature is also decreased. The cause of this effect has been traced back to a sort of partial fusion occurring at the point of contact, due to

local heating by the current. As a result, the whisker adheres to the crystal and the area of contact increases. The whisker, then, no longer acts as a mere point of easy contact, and so rectification disappears. Lowering the temperature or the voltage will not, in general, cause a detachment of the whisker from the surface of the crystal, and thus the former condition persists.

In the table below we give a typical result for D.C. rectification at three temperatures (see also figure 1). I_1 is the current from the crystal to the point, while I_2 is the current in the opposite sense. N is the rectification ratio, and is equal to the difference $(I_1 - I_2)$ divided by the larger current. The unit in which the currents are expressed is 0·04 ma.

D.C. rectification by bornite at different temperatures

Voltage (V)	At 30° c.			At 120° c.			At 170° c.		
	I_1	I_2	N	I_1	I_2	N	I_1	I_2	N
0·40	41	8·5	0·80	22	25	−0·12	35	46	−0·24
0·32	32·5	7·5	0·77	16	19	−0·16	25	34	−0·26
0·24	23·5	6·5	0·74	10·5	11·5	−0·09	17	23	−0·26

It will be observed that I_1 changes irregularly with temperature, first increasing and then decreasing. I_2 usually increases with rise of temperature, but not in accordance with the equation which was found by Waterman (1923) for the electric conductivity of molybdenite.

§ 4. DISCUSSION OF THE RESULTS

The reversal of rectification by heat appears to be of a thermoelectric nature. We can represent the character of the crystal by the character of a thermocouple whose cold junction is gradually raised to higher temperatures. It will be found

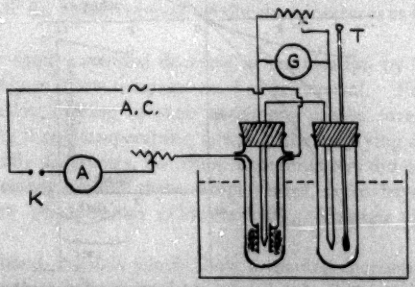


Figure 2. Thermoelectric inversion by raising the temperature of the cold junction

that, at a certain critical temperature of cold junction, any excess in the temperature of the other junction will cause reversal in the direction of the thermo-e.m.f This can be clearly illustrated by means of an experiment arranged as in figure 2.

The two junctions of a copper-steel thermoelement are placed within two tubes immersed in a wide-range thermostatic bath. The temperature excess in one junction is produced by means of a heating coil wound over a short capillary tube of pyrex glass enclosing the junction. The coil is wrapped in asbestos to reduce heat losses. It is carefully insulated from the thermocouple, and is preferably heated by alternating current in order to eliminate leakage of electricity to the thermocouple circuit. The difference between the temperatures of the two junctions is roughly proportional to the square of the heating current. At a fixed temperature of the thermostat, different heating currents are sent through the coil, and the deflection of a shunted galvanometer due to the thermo-e.m.f. produced is observed. The results are represented in figures 3, 4 and 5.

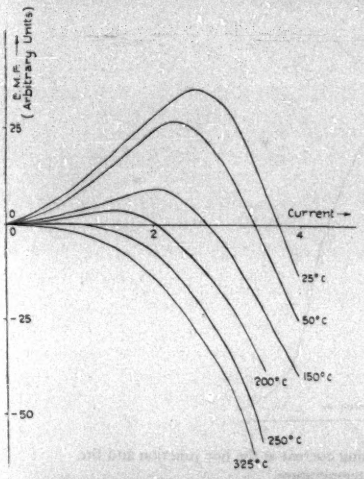


Figure 3. Relation between the heating current at the hot junction and the resulting e.m.f. for different temperatures of the cold junction.

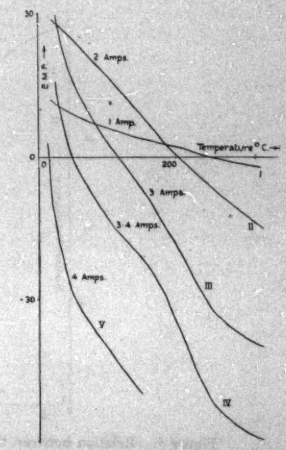


Figure 4. Relation between the temperature of the cold junction and the e.m.f. due to a fixed heating current at the other junction.

From curve I (figure 4) it will be seen that raising the temperature of the cold junction above 250° c. will cause reversal in the direction of the deflection, while from curve V it will be observed that lowering the temperature of the cold junction below 16° c. will produce reversal. In other words, reversal occurs either by increasing or by decreasing the temperature of the cold junction, according to the value of the current that produces the excess of heating in the hot junction. As we increase the temperature of the cold junction we decrease the excess that causes reversal. From figure 5 we deduce that the temperature

at which reversal occurs for any excess of heating in the hot junction lies below 266° c., which is nearly the neutral temperature of the thermocouple used.

The analogy between the behaviour of a rectifying crystal and that of a thermocouple is clear. We may represent the cold junction by the whole crystal, and the hot junction by the point of contact (assuming the Joule effect to be the prevailing one). The heating current may be represented by the effective current in the crystal (in A.C. measurements), and the thermo-e.m.f. by the rectified current. For small applied voltages we should expect reversal to occur at low or at high temperatures, according as the contact has a low or a high neutral point. The curves of figure 4 are analogous to the temperature-rectification curves. The

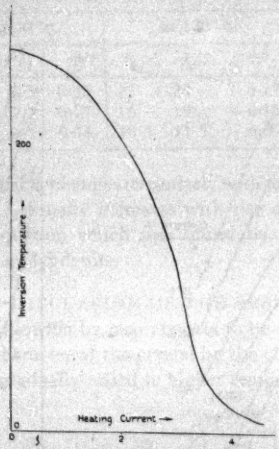


Figure 5. Relation between the heating current at the hot junction and the inversion temperature.

analogy suggests that rectification either increases or decreases with rise of temperature according to the working conditions, as is experimentally found.

In molybdenite the authors found that the inversion temperature is lower the higher the applied voltage. This result is in accordance with the curve of figure 5. In bornite, however, this is not always the case. The discrepancy is more than can be accounted for by experimental error.

§5. CONCLUSION

The results of our experiments show that rectifying crystals, which are sensitive to heat variations, reverse rectification when the temperature is changed in the direction that favours diminution in the rectification. These crystals are known to have abnormally large thermoelectric forces against metals (Wick, 1907;

Pierce, 1909). Their behaviour towards heat indicates that their rectifying action may be attributed, at least in part, to a thermoelectric effect. This point, in general, supports the thermoelectric theory of rectification developed by Eccles (1910 and 1913).

REFERENCES

Eccles, W. H., 1910. *Proc. Phys. Soc.* **22**, 360 and 477 ; 1913. *Ibid.* **25**, 4 and 273.
Flowers, A., 1909. *Phys. Rev.* **29**, 455.
Khastgir, S. and Das Gupta, A., 1935. *Phil. Mag.* **19**, 557.
Pierce, G., 1909. *Phys. Rev.* **28**, 175.
El Sherbini, M. A. and Yousef, Y. L., 1939. *Phys. Soc. Proc.* **51**, 449.
Waterman, A. T., 1923. *Phys. Rev.* **21**, 541.
Wick, F. G., 1907. *Phys. Rev.* **25**, 390.

MEASUREMENT OF TENSION IN LIQUIDS BY MEANS OF A METAL BELLOWS

By R. S. VINCENT, M.Sc., A.M.I.E.E.

Communicated by A. Simons, M.Sc. *Received* 29 *November* 1940

ABSTRACT. The paper commences with an historical review of published work. A form of apparatus is then described in which a metal bellows, closed by a needle valve, contains the liquid under test, and the possibility of inward leakage of atmospheric air is avoided by completely surrounding the bellows and valve by the same liquid. A gradually increasing tensile load is applied to the bellows until the liquid breaks. The hydrostatic tension which existed in the liquid immediately prior to breaking is then calculated from the load and the known effective cross-section of the bellows.

The liquids so far examined by this method are ethyl alcohol, ethyl ether and mineral oil. The scope of the paper has been confined to results obtained in the temperature range 0° c. to 30° c.

§1. INTRODUCTION

One of the most interesting properties of liquids is their ability to withstand hydrostatic tension. The best known example of this property is the fact that mercury completely filling a barometer tube can be raised to a level considerably above the barometric height (Reynolds, 1878). This phenomenon was explained by Laplace as being due to the adhesion of the mercury to the glass, and to its own power to withstand tension (Dixon, 1909–10).

The earliest work on the actual measurement of tension in liquids appears to have been described by F. M. Donny (1843), who found that it was possible for a column of sulphuric acid 1·255 metres long to hang in a vertical glass tube sealed at the upper end, even when the atmospheric pressure was removed from the lower end. Donny pointed out that a bubble, sufficiently small to have surface-tension forces capable of supporting the hydrostatic head of the liquid below, will not destroy the tensile state.

In 1850 Berthelot showed that water can resist a considerable tension. He filled with water a strong capillary tube sealed at one end and drawn to a fine point at the other. The tube was then cooled, thus drawing a small volume of air into the drawn-out point, which was then sealed off. The tube was then heated, causing the air to dissolve and the water to fill the tube completely. Upon re-cooling it was noticed that the water continued to occupy the whole volume of the tube at a temperature lower than that at which the water had first filled the tube. Finally, the water column broke, leaving a space at the end of the tube. Berthelot found that the water had increased in volume by one part in 420.

By assuming that the extensibility of the water under tension was the same as its compressibility, he deduced that the observed extension would correspond to a tension of about fifty atmospheres. Berthelot also made similar experiments with other liquids and solutions. In one of these experiments he estimated a tension of 100 atmospheres in sugar solution (Tait, 1890).

In 1878 Osborne Reynolds described a centrifugal method of applying tension to a liquid column (Reynolds, 1882). On a lathe he rotated a sealed U-tube, the short arm of which was completely filled with water, the other arm being occupied partly by water and partly by water vapour. By observing the maximum speed of rotation possible without causing the liquid to break, the maximum tension sustained by the liquid was calculated. Osborne Reynolds obtained a value of 4·8 atmospheres for water.

In 1892 Worthington described further experiments which he carried out using the centrifugal method of Osborne Reynolds, and obtained a value of 7·9 atmospheres for ethyl alcohol. Worthington then described a method he had devised which enabled both the extension and the tension of ethyl alcohol to be measured. The tension in the liquid was produced by cooling as in the case of Berthelot's method, but the tension was measured directly by means of a large ellipsoidal bulb containing mercury. This was enclosed within the glass apparatus, and was provided with a capillary tube in which the top of the mercury column indicated changes of volume of the bulb. This tension-measuring device was previously calibrated with positive pressure, and the scale extended in the negative direction by extrapolation. The greatest tension estimated by Worthington was 17 atmospheres at a temperature of about 16° c. Worthington also found quite high tensions to be developed when a small piece of copper foil was enclosed within the apparatus. Under tension this showed a tendency to adhere to the glass, and the break always occurred at the point of contact.

In 1895 Dixon and Joly carried out some experiments to find out if water adhered to the conducting tubes of plants sufficiently vigorously to transmit the stress needed to raise the transpiration stream. Berthelot's method was employed, a small piece of wood being inserted in the tube. It was found that the break almost invariably occurred on the surface of the glass, not on the surface of the enclosed wood.

In 1909 Dixon described some further experiments on these lines in which he deliberately left a trace of air before sealing the tube. A small piece of yew wood was enclosed in one of the tubes and in another some minute copper shavings. It was found that the liquid column always broke with a metallic click at tensions estimated at values up to 150 atmospheres.

Askenasey in 1895 performed an ingenious experiment to illustrate the theory of transpiration in plants. A barometer tube containing gypsum solution was terminated at its upper end by a funnel sealed by means of a tightly fitting plate of gypsum, the lower end of the tube dipping into a vessel of mercury. As the solution was absorbed by the gypsum a mercury column was drawn up

which exceeded the barometric height by 13 cm.; hence a tension corresponding to 13 cm. of mercury was produced in the solution. Hulett (1903) later made similar experiments using a porous clay disc treated with potassium ferrocyanide and copper sulphate. He obtained a tension in water corresponding to a head of 37·7 cm. of mercury.

In a valuable paper by Julius Meyer, published in 1911, a comprehensive review of previous work is given together with an account of Meyer's own measurements of tension in ethyl ether, ethyl alcohol and water. Meyer's results are quoted in the *International Critical Tables* (Young and Harkins, 1926). The maximum values of tension observed by Meyer were 34 atmospheres in water at 24°·4 c., 39 atmospheres in alcohol at 22°·5 c., and 72 atmospheres in ether at 17°·7 c. The method used was similar in principle to that of Worthington, the tension being produced by cooling the liquid within a glass vessel. In Meyer's apparatus, however, the tension-measuring device was the glass spiral manometer of Ladenburg and Lehmann (1905) as modified by Abegg and Johnson (1908). This device consists of a helical glass tube upon which a mirror is mounted. The slight rotation of the end of the glass helix due to the internal tension may be observed by means of the attached mirror. The device must be previously calibrated by means of a positive pressure, and the range extended in a negative direction by extrapolation.

In 1912, H. M. Budgett showed that the pull necessary to rupture a film of various liquids between two flat steel surfaces under vacuum is sometimes as much as about four atmospheres. He considered this to be much lower than the maximum tension which the liquids themselves were capable of withstanding.

It will be noted in the above review of published literature that there are wide variations between the values of maximum tension given by the various experimenters. Published results may therefore be said to be unsatisfactory in the sense that no two authorities are in agreement as regards the maximum tension it is possible to apply to liquids.

After consideration of the various methods which have been employed, the writer formed the opinion that uncertainty may have arisen in some cases in calculating the tension from the observed results. For example, Worthington and Meyer used their tension-measuring devices in ranges over which they were not actually calibrated, although it is difficult to believe that large errors could have arisen in that way. Again, Berthelot and Dixon calculated the tension on the assumption that the elasticity of the liquid under tension has the same value as its compressibility, even up to the maximum value of tension sustained, and in spite of the possible existence of microscopic stable bubbles within the stressed liquid. Thus the only two methods which are theoretically irreproachable, those of Osborne Reynolds and Budgett, give comparatively low results. In fact it will be shown that the results obtained depend chiefly upon the choice of method. Further comments on the various methods will be made later in the present paper when considering the results of the experiments to be described.

As a first step to discovering the reasons for different values of the maximum tension being obtained by different methods, the experiments described in the present paper have been carried out, and proposals for the further elucidation of the problem have been made. The simple and direct method described below has the advantage that the applied tension is known with a high degree of certainty, and that the temperature may be kept constant and uniform at any required value. A further advantage is the ease with which any bubbles formed can be eliminated. The impossibility of ejecting bubbles once formed is a serious limitation in a permanently sealed apparatus..

§2. DESCRIPTION OF APPARATUS

The apparatus is illustrated in figures 1 and 2. An important feature of the design is that the bellows and the needle valve which is used to make the final seal are completely immersed in the liquid that is being tested, thus avoiding the possibility of leakage from the atmosphere. The mechanical connection between the lower end of the bellows and the applied load has been brought out of the enclosing glass vessel through a mercury barometer. The cylindrical cup containing mercury, which seals the lower end of the barometer tube, is screwed on to the steel rod, through which the tensile force is applied, and moves up and down with it. The needle valve may be opened or closed as required by rotating the vertical steel rod by means of the fitting which is attached at its lower end for suspending the adjustable load. The form adopted for this adjustable load has been the subject of some preliminary experiments. It was found that the use of a scale pan to which weights were added proved unsatisfactory, since the sudden break of the liquid always occurred during alteration of the weights on the pan. This led to some uncertainty in the result, so that a method of increasing the load continuously without jerking the apparatus had to be devised. This consisted of suspending a cylindrical can in place of the scale pan and slowly filling it with water from a small jet. The can was provided with a steel-wire handle which could be attached to the fitting at the bottom of the vertical rod by means of a hook. A set of lead weights was provided, which could be adjusted in 500-gramme stages. The normal procedure was to place some of these weights in the can, then increase the weight gradually by slowly adding water through the glass jet until a break in the liquid occurred. The glass jet was supplied from the water main through a constant-head device, and gave a flow of approximately 60 grammes of water per minute.

A copper tube through the rubber stopper at the top of the apparatus supports the screw-thread of the needle valve to which it is brazed, and also provides a connection to the interior of the glass enclosure. For this purpose, several holes are drilled through the tube just below the rubber stopper. This tube serves for admission of the liquid to be tested, and provides a means of evacuating the interior of the apparatus.

Just below the point where the steel rod is soldered into a brass plug in the

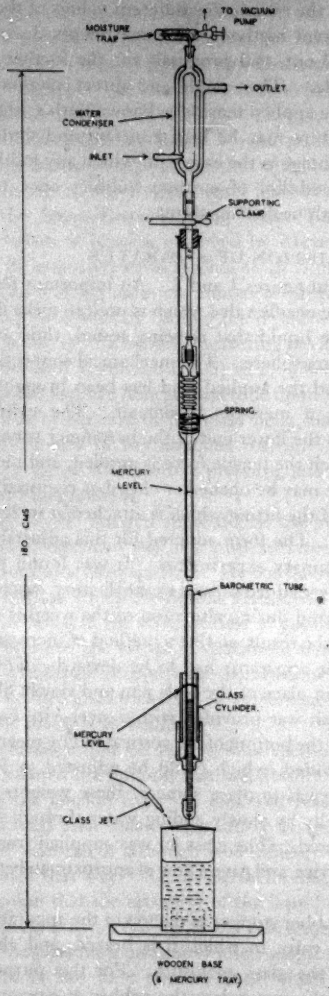


Figure 1.

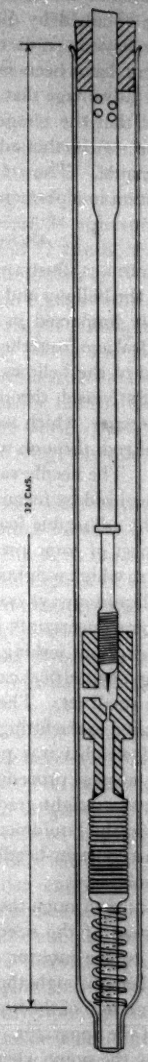


Figure 2.

lower end of the bellows, a small helical spring is provided which is free to move on the rod. This spring normally rests on the shoulders at the top of the barometer tube and serves as a buffer to prevent fracture of the glass while handling the apparatus. A water-cooled condenser is connected to the top of the glass enclosure, and above this is a moisture trap containing glass wool and phosphorus pentoxide.

In the process of filling the apparatus with liquid, application of heat is required for driving off residual air. This heating was done partly by applying a gas flame carefully to the outside of the glass enclosure, but to ensure that sufficient heat was reaching the bellows, an electrical method of heating was also employed. A solenoid was placed over the apparatus and an alternating current passed through it from a suitable tapping of a transformer. Heat was thus generated in the top end of the steel rod ; this ensured that in the case of a volatile liquid, boiling actually took place within the bellows and was not limited to the region just inside the interior surface of the glass. This easily controlled form of heating was found to be very useful throughout the work. In the experiments on ether the upper part of the apparatus was enclosed in a thermostatically controlled chamber the temperature of which could be adjusted to the required values in the range above atmospheric temperature. For the temperature of 0° c. a metal container surrounding the bellows region was filled with melting ice.

Throughout the present work the temperature was measured by means of a mercury thermometer hanging close to the glass enclosure near the bellows.

§ 3. DETERMINATION OF CROSS-SECTIONAL AREA OF BELLOWS

In calculating the hydrostatic tension produced by a given load on the bellows, it was at first sight uncertain which diameter of the bellows should be used in the calculations. It was therefore decided to determine the effective cross-sectional area of the bellows by a direct experiment. A bellows assembly, as shown in figure 2, was provided with two pointers, one attached to the top of the bellows and the other to the bottom. The distance between these two pointers could be read by means of a travelling microscope. First, the distance between the two pointers was measured with atmospheric pressure both inside and outside the bellows. The interior of the bellows was then evacuated down to a negligible pressure. A scale pan was then attached to the lower end of the bellows and the total weight was gradually increased until the original distance between the pointers was obtained. The load required was 1280 grammes, which was equal to the existing atmospheric pressure acting over an area of 1·25 sq. cm. The mean internal diameter of the bellows was 0·952 cm. and the mean external diameter 1·460 cm., so that the effective area of the bellows did not differ greatly from the mean of the two calculated areas, which was 1·19 sq. cm.; the experimentally determined effective area, however, has been used in all calculations. With the bellows assembly set up as described, and the valve open, the distance between the pointers was measured for various loads. The results

showed that the extension was proportional to the load for extensions up to at least 2 mm., and in this range no permanent set could be detected. The force required per millimetre extension was approximately 400 grammes weight.

§4. DESCRIPTION OF EXPERIMENTS ON ETHYL ALCOHOL AND THE RESULTS OBTAINED

After filling the apparatus shown in figure 1 by pouring in absolute alcohol to the required level, the alcohol within the apparatus was boiled by applying a gas flame, and to ensure that boiling really occurred inside the metal bellows, the heating was augmented by means of the solenoid placed around the apparatus at the region to be heated. After continuing this boiling for a few minutes, the apparatus was allowed to cool down to room temperature. Although there was air over the surface of the alcohol during cooling, it did not apparently penetrate into the bellows via the alcohol through the small passage in the needle valve. The needle valve was then closed by rotating the vertical steel rod from its lower end. The can was then attached to its hook, and the flow of water from the glass jet was started. Close attention was paid to any visible movement of the can, the observer being ready to turn off the water flow immediately a sudden downward movement of the can on to its supporting base was observed. The wooden base was placed below the can at a distance of about 2 mm., and in this way the movement of the bellows upon rupture of the liquid was limited to a safe amount. It is obvious that the bellows could not, without permanent injury, support the whole load when the liquid column had broken.

After a break in the liquid had been produced and the jet turned off, the can was removed and weighed complete with contents. Upon opening the needle valve immediately after a break and compressing the bellows, a bubble was nearly always seen to escape from the valve opening.

The following example will serve to illustrate the method of calculation. The measurements were obtained on absolute alcohol inserted in the apparatus by boiling under atmospheric pressure for about 10 minutes, followed, after cooling, by boiling under reduced pressure.

Temperature	5° c.
Barometric pressure	765 mm. Hg.
Pressure in apparatus	Atmospheric
Breaking load	4375 grammes
Effective cross-section of bellows	1·25 sq. cm.
Upthrust due to atmospheric pressure on the bellows $76{\cdot}5 \times 13{\cdot}6 \times 1{\cdot}25$	$=1300$ grammes
Actual load on liquid therefore	$=4375-1300$ $=3075$ grammes
Therefore tension in liquid prior to break	$=\dfrac{3075}{1{\cdot}25 \times 76 \times 13{\cdot}6}$ $=2{\cdot}38$ atmospheres

No attempt has been made in the present work to correct for the small force exerted by the bellows due to any slight extension prior to the rupture of the liquid.

A summary of the results obtained on ethyl alcohol is given in the table below. All the results quoted were obtained with air at atmospheric pressure over the surface of the alcohol in the enclosure. Measurements at 0° c. were obtained by placing melting ice in the jacket around that part of the apparatus containing the bellows.

Table 1. Summary of results on ethyl alcohol

Treatment prior to test	Temp. (° c.)	Tension (atmos.)
Boiled at atmospheric pressure, left 18 hrs.?	6	2·09
Bubble escaped, left 30 min.	7	1·99
Bubble escaped, left 7 hrs.	9	2·05
Vacuum applied momentarily, left 35 min.	11	2·12
Bubble escaped, vacuum applied momentarily, left 16 hrs.	5	2·38
Bubble escaped, left 1 hr.	8	2·27
Bubble escaped, left 30 min.	10	2·27
Bubble escaped, left 15 min.	10	2·04
Bubble escaped, vacuum applied momentarily, left 6 hrs.	12	1·34
Bubble escaped, left 16 hrs.	0	1·88
Bubble escaped, left 5 hrs.	0	1·25

The periods for which the liquid was left prior to testing were spent at atmospheric pressure with the needle valve slightly open. It will be seen that even the highest value obtained (2·38 atmospheres) is very much lower than the values of 17 atmospheres obtained by Worthington (1892), and 39 atmospheres obtained by Meyer (1911). This seems to suggest that the maximum tension is governed by the degree of success with which residual gas has been removed from the liquid, and also by freedom of the enclosure from occluded gases. Glass is likely to be better than metal in this respect. The maximum tension which liquids in contact with metal surfaces will in practice sustain is, however, of considerable interest.

§ 5. EXPERIMENTS ON ETHYL ETHER

For these experiments the top half of the apparatus was enclosed in a thermostatically controlled chamber having a double glass front through which observations could be made.

Using the same procedure as already described for ethyl alcohol, the following results have been obtained for ether. Only the highest results which were obtained at each temperature are quoted in the table below:—

Table 2. Measurements of tension in ethyl ether

Temp. (° c.)	Treatment	Maximum tension (atmos.)	No. of tests
0	Boiled, cooled, boiled, cooled	2·21	19
19·2	Boiled 55 min., cooled	0·69	12
29·9	Boiled 100 min., cooled	0·00	10

The value 0·00 for 29·9° c. was the actual result of the calculation, and was also the highest of the ten results obtained, since all the others were negative. A negative result simply means that a positive pressure must be maintained on the liquid to prevent the formation of a bubble.

A very interesting series of results was obtained at 0° c., which are worth reproducing in detail. After boiling the ether for one hour, ice was applied. One hour later the results plotted in figure 3 were obtained. Any bubbles formed were eliminated from the bellows after each break. The tendency of the maximum tensions to increase after successive breaks is clearly indicated in the graph. The occasional low results, which have been found to occur throughout the present work, may be due to sudden evolution of a bubble of occluded gas from the enclosure, or possibly to the coalescence of two existing bubbles. A time lag has been noted throughout the work, not only for ether but also for the other liquids, which may be accounted for by delay in the evolution of a bubble from the

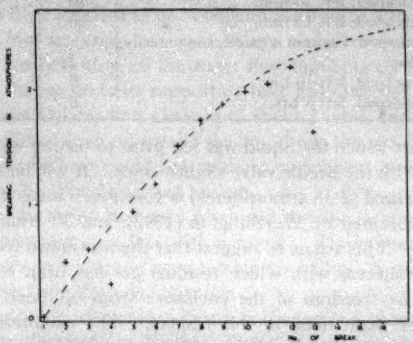


Figure 3. Graph illustrating successive breaks in ether.

enclosure, or by the time required for coalescence of a number of smaller bubbles. In the course of preliminary tests to determine if the liquid was sufficiently "degasified" to withstand tension, it was frequently noticed that a particular load would be supported for a few seconds before the liquid broke; in other words, there was a time lag between the application of the load and the breaking of the liquid.

§ 6. EXPERIMENTS ON MINERAL OIL

The mineral oil employed in the following experiments was of the paraffin-base type, and described by the makers as triple filtered. At 20° c. it had a viscosity of 16·5 poise and a density of 0·893 gm. per c.c. For the purpose of filling, the condenser and trap were removed from the apparatus. A glass globe having a ground-glass plug cock at the top and an outlet tube at the bottom closed by a short piece of rubber tube and clip was half filled with the oil. The oil was then

"degasified" by applying a vacuum, via the plug cock, by means of a rotary oil pump, and at the same time heating the globe with a gas flame and agitating it from time to time to assist gas removal. Although this process was continued for some hours, the stage was not reached when it was impossible to liberate another bubble by shaking. The surface of the hot oil was, however, quite placid in the absence of agitation, with the pressure reduced to approximately 0·1 mm. Hg.

After leaving the apparatus evacuating all night, the oil globe was attached by means of a T-piece. This was evacuated to remove entrapped air in the connection. By the appropriate operation of clips, the required quantity of degasified oil was admitted to the apparatus. The absolute pressure in the apparatus was kept down to approximately 0·1 mm. Hg., and the oil surrounding the bellows was heated with a Bunsen flame to augment the heat generated by the solenoid in the top end of the steel rod. Under this treatment it appeared that the oil would continue to produce bubbles indefinitely. It was therefore decided to discontinue degasification, and to apply atmospheric pressure by admitting atmospheric air over the surface of the oil. After cooling, the needle valve was closed. The first tensile test gave a total load of 2594 grammes, which corresponded to a tension of 1·01 atmospheres.

Vacuum and heat were again applied as before and continued for three hours down to a final pressure of 0·1 mm. Hg. Even after this treatment it was still possible to produce bubbles by expanding and contracting the bellows. Atmospheric air was admitted over the oil surface and, after cooling, the valve was closed. This time a tension of 1·47 atmospheres was measured. This higher value was attributed to improved degasification. Heat and vacuum treatment was again carried out as before, but apparently with greater success, because it was found possible to fill the bellows by drawing in the liquid under vacuum by means of its own tension. This procedure made it unnecessary to admit atmospheric air over the oil for the purpose of filling the bellows. After cooling, the valve was closed and a pressure equal to that of one atmosphere was applied to the liquid inside the bellows by means of a laboratory balance which was arranged to push the steel rod upward with a force of 1300 grammes weight. A tensile test with the apparatus still under vacuum gave a definite break at a load of 2866 grammes, which corresponded to 2·22 atmospheres.

Even the few results so far described indicate that improved degasification increased the power of the liquid to withstand tension.

Further work was directed to determining the effect of even better degasification. A geared electric motor was set up so as to lower a 500-gramme weight on to the pan of the laboratory balance previously mentioned, so that an upward force of 500 grammes was intermittently applied to the bottom of the bellows at about 5-second intervals. With the oil kept hot by means of the solenoid, a small bubble was seen to escape from the bellows at every cycle. The evolved gas was collected during a period of 15 minutes; it was found to be completely absorbed

by phosphorus pentoxide. It is therefore probable that it consisted entirely of water vapour.

After the above treatment, involving repeated application of pressure and tension, the oil appeared to be in the best state of degasification so far achieved. After 18 hours under a pressure of one atmosphere, a tension of 1·82 atmospheres was observed.

It was decided to carry out further tests on this same sample without opening the bellows. The pressure of one atmosphere was therefore re-applied by means of the balance. After a further 24 hours, that is 42 hours from the enclosure of the sample, the value of 2·45 atmospheres was obtained. The pressure of one atmosphere was again applied, and at 66 hours after sealing the tension of 2·94 atmospheres was measured. Further readings were taken on this sample in a similar manner, all of which are quoted in table 3 below, together with a summary of the results so far discussed. The low final reading may possibly be explained by the gradual accumulation of chemically generated gas or moisture, or perhaps by the liberation of occluded gas from the enclosure.

Table 3. Summary of results on mineral oil

Temp. (°c.)	Hours since last break	Hours since sealing	Tension (atmos.)
13	18 (since start)	—	1·01
6	21	—	1·47
10	48	—	2·22
7	18	—	1·82
10	24	42	2·45
9	24	66	2·94
12	24	90	2·59
13	24	114	2·72
12	24	138	2·51
9	24	162	2·82
8	24	186	$<0{\cdot}6$

It is concluded from the above results that the application of hydrostatic tension is an effective method of carrying out the final stages of degasification, and that the maximum tension which the liquid will withstand appears to be a good indication of the success of the degasification process.

§7. DISCUSSION OF PRESENT AND PREVIOUSLY PUBLISHED RESULTS

The mutual attraction between the molecules of a liquid results in an "intrinsic pressure" within the liquid. For liquids at ordinary temperatures the intrinsic pressure has a very high value. For water it is $1{\cdot}06 \times 10^{10}$ dynes per sq. cm., i.e. about 10,400 atmospheres. It is generally assumed that in a perfectly gas-free liquid this intrinsic pressure would have to be overcome to

pull apart a column of liquid (Champion and Davy, 1936). In practice, liquids have never been found to sustain tensions of anything like this magnitude. It is interesting to note in passing that neither do solids exhibit their theoretical strength—for example, the mechanical tensile-strength of rock crystal as calculated from electrical theory is many hundreds of times that actually observed. The discrepancy may be explained by supposing that the rupture of the crystal is not produced simultaneously across the whole cross-section (Joffé, 1928). This divergence between theory and practice has been pointed out by Professor R. Fürth (Born and Fürth, 1940). It is easy to see that such a result might arise in solids due to inhomogeneity of structure which would cause non-uniformity of strength and of stress. As the writer has previously suggested (Vincent, 1940), liquids might be expected to possess certain advantages over solids for the study of cohesion problems in view of their homogeneity and the uniformity of the applied stress. The difference between the intrinsic pressure and the actual tension which liquids will in practice sustain may be due to the difficulty of rendering a liquid perfectly gas-free, but the more likely explanation is that the mechanism of breaking is not the one implied by consideration of the intrinsic pressure. The writer's view, based on the experiments herein described, and on previously published work, is that the tension in a liquid is in practice always relieved by the formation of a bubble sufficiently large to be unstable under the existing tension. The metallic click reported in some experiments gives rise to the idea of a fracture of the liquid, but this may equally well be due to the rapidity of expansion of the unstable bubble, resulting in a very sudden relief of the tension. Further, it is not justifiable to assume that a perfectly gas-free liquid would behave differently, since bubbles in a liquid under tension could be vacuous or contain only the vapour of the liquid, gas not being necessary to ensure their expansion, and they might easily originate as a result of the random thermal agitation of the liquid molecules. Professor R. Fürth has suggested that holes do in fact exist in liquids as part of their normal constitution (Fürth, 1940).

It may be that the mechanism which results in the sudden relief of hydrostatic tension in a liquid has a lot in common with the rupture of a solid under a tensile load, particularly while the solid is yielding under the applied load. The results of Professor M. Born and Professor R. Fürth, indicating the close relationship between breaking and melting, are of interest in this connection (Born and Fürth, 1940).

Returning now to the results of the present paper, it is interesting to calculate the maximum size of bubble which would withstand the observed maximum tensions of 2·38 atmospheres in alcohol, 2·21 atmospheres in ether, and 2·94 atmospheres in mineral oil. In the case of alcohol, its vapour pressure would be exerted inside the bubble; the pressure due to other gases, however, will be assumed to be negligible.

Now, if S is the surface tension, the maximum hydrostatic tension T which the bubble will withstand is given by

$$T = \frac{2S}{R} - p \qquad \ldots\ldots(1)$$

or

$$R = \frac{2S}{T+p}, \qquad \ldots\ldots(2)$$

where in this case p is the saturated vapour pressure of alcohol.

Substituting values,

$$R = \frac{2 \times 23{\cdot}1}{2{\cdot}38 \times 1{\cdot}013 \times 10^6 + \frac{17{\cdot}7}{750} \times 10^6}$$

$$= 1{\cdot}9 \times 10^{-5} \text{ cm.}$$

The highest tension in ether was observed at 0° c., at which temperature the saturated vapour pressure is 184 mm. Hg. and the surface tension 18·8 dynes/cm. We have, therefore,

$$R = \frac{2 \times 18{\cdot}8}{2{\cdot}21 \times 10^6 \times 1{\cdot}013 + \frac{184}{760} \times 10^6}$$

$$= 1{\cdot}51 \times 10^{-5} \text{ cm.}$$

In the case of the mineral oil at 9° c., the surface tension is approximately 30 dynes/cm. and the saturated vapour pressure is negligible, so that

$$R = \frac{2 \times 30}{2{\cdot}94 \times 1{\cdot}013 \times 10^6} = 2{\cdot}0 \times 10^{-5} \text{ cm.}$$

The above calculations indicate the radius of the largest bubble which in each case could exist under the maximum tension observed. An interesting point which arises out of the extreme smallness of these bubbles is the fact that the bubbles which escaped from the bellows opening immediately after a break were usually between 1 mm. and 3 mm. in diameter, even at atmospheric pressure. It seems, therefore, that upon the sudden expansion of the minute bubbles, a considerable amount of gas must be liberated from the liquid into the bubble; alternatively, it must be assumed that an appreciable quantity of gas is suddenly liberated from the enclosure in the form of a bubble, which was not present prior to the break.

It is clear that the repeated breaking of the liquid by application of hydrostatic tension is an effective method of degasifying both the liquid and its enclosure, and that the value of the maximum tension is a good indication of the success of the degasification process. The fact that the bubble size has worked out to be of the same order in each case may be merely coincidence, or there may be some reason for it. Although one cannot assume that a particular bubble size will always be reached, it is interesting to consider equation (1) on the assumption

of a constant bubble size. This relation now indicates in a general way the effect of temperature on the maximum tension. It is clear that, due to the rapid increase of vapour pressure, the maximum tension would fall at an increasing rate as the temperature is raised. This would be worthy of further investigation, but the suggestion seems to be borne out by the three results for ether. The formula may perhaps also indicate in a general way what type of liquid would withstand the greatest tension for a given standard of degasification; apparently one should choose a liquid of low vapour-pressure and high surface-tension. This is another question which would merit investigation.

The results given in the present paper, together with published results, are summarized in the table below :—

Table 4. Maximum tensions in liquids (at laboratory temperatures unless otherwise stated)

Liquid	Method	Tension	Observer.
Water	Berthelot	50	Berthelot
	,,	150	Dixon
	,,	157	Vincent
	Reynolds	4·93	Reynolds
	Meyer	34	Meyer
Alcohol	Reynolds	7·9	Worthington
	Worthington	17	,,
	Meyer	39	Meyer
	Vincent	2·38	Vincent
Ether	Meyer	72	Meyer
	Vincent	2·21 (0° c.)	Vincent
Mineral oil	,,	2·94	,,
Sulphuric acid	Reynolds	11·8	Worthington
Various liquids	Budgett	4 (max.)	Budgett

It appears from the above table that the result obtained depends chiefly on the choice of method. The bellows method, and the methods of Osborne Reynolds and Budgett, in which it seems probable that the tensions are known with the greatest certainty, give comparatively low results.

The method of deducing the tension at the moment of break employed in Berthelot's method will give a falsely high answer if stable bubbles exist in the liquid prior to the break. This would mean that part of the extension is only apparent, hence the tension calculated from it would be fictitious. It is possible, however, that the difference in the results is due merely to the different degrees of success achieved in the removal of gas from the liquid and its enclosure.

From what has been suggested previously upon the relief of tension resulting probably from bubble instability, it seems that there is probably a much lower limit to the tension which a liquid can sustain than has previously been thought, i.e., much lower than the intrinsic pressure of the liquid. The first step to evaluating this limit would presumably be to determine if the results obtained

by Berthelot's method are true. There is no doubt that the method does yield high results. With a little care the writer has obtained a value of 157 atmospheres by this method, the temperature at which the tube filled being 69° C., and that at which the column broke with a click 54° C., corresponding to an extension of 0·78 per cent. To make the method reliable, some independent method of measuring the tension within the stressed liquid must be found. The writer hopes to describe some experiments in this direction in due course; in the meantime it is hoped that a contribution has been made to the knowledge of this interesting subject.

§8. ACKNOWLEDGEMENT

For granting permission to publish this paper, the author wishes to thank Callender's Cable and Construction Company, in whose laboratories the experimental work was carried out.

REFERENCES

ABEGG, R. and JOHNSON, G., 1908. *Z. phys. Chem.* **61**, 457.
ASKENASEY, 1895. *Verh. Naturh. Med. Ver. Heidelberg.*
BERTHELOT, 1850. *Ann. Chim.* (*Phys.*), **30**, 232.
BORN, M. and FÜRTH, R., 1940. *Nature, Lond.*, **145**, 741.
BUDGETT, H. M., 1912. *Proc. Roy. Soc.* A, **86**.
CHAMPION and DAVY, 1936. *Properties of Matter* (Blackie and Son, Ltd.).
DIXON, 1909. *Sci. Proc. R. Dublin Soc.* **12**.
DIXON and JOLY, 1895. *Philos. Trans.* **186**, 563.
DONNY, F., 1843. *Acad. Bruxelles*, **17**; 1846. *Ann. Chim.* (*Phys.*), **16**, 167.
EDSER, 1911. *General Physics*, pp. 281, 342, 356, 557.
FÜRTH, R., 1940. *Proc. Phys. Soc.* **52**, 768.
HULETT, G. A., 1903. *Z. phys. Chem.* **42**.
JOFFÉ, A. F., 1928. *Trans. Faraday Soc.* **24**. (General discussion on Cohesion and related problems.)
KAYE and LABY, 1928. *Tables of Physical and Chemical Constants* (seventh edition), p. 41.
LADENBURG, E. and LEHMANN, E., 1905. *Ber. dtsch. phys. Ges.* **4**, 20.
MEYER, JULIUS, 1911. *Zur Kentniss des Negativen Druckes in Flussigkeiten.* (Halle a. d.S: Knapp.) (*Abhandlungen d. Deutchen Bunsen-Gesellschaft*, Nr. 6.)
POYNTING and THOMSON, 1902. *Properties of Matter*, p. 122.
REYNOLDS, OSBORNE, 1878, *Mem. Manchr Lit. Phil. Soc.* **17**, 159.
REYNOLDS, OSBORNE, 1882. *Mem. Manchr Lit. Phil. Soc.* **7**, 1. (Read 16 April 1878.)
TAIT, P. G., 1890. *Properties of Matter* (A. and C. Black), 200.
VINCENT, R. S. 1940. *Nature, Lond.*, **145**, 970.
WORTHINGTON, A. M., 1892. *Philos. Trans.* A, **183**, 355 ; 1890. *Proc. Phys. Soc.*, and *Nature, Lond.*, **42**, 261.
YOUNG, T. F and HARKINS, D., 1926. *International Critical Tables*, **4**, 434.

ON ULTRA-HIGH-FREQUENCY OSCILLATIONS GENERATED BY MEANS OF A DEMOUNTABLE THERMIONIC TUBE HAVING ELECTRODES OF PLANE FORM

By W. A. LEYSHON, Ph.D., F.Inst.P.,
London (R.F.H.) School of Medicine for Women

Received 30 *October* 1940

ABSTRACT. An account is given of experimental results obtained with a demountable tube shown to the Physical Society in November 1937.

The symmetrical double kathode, double positive-grid arrangement shows distinct advantages over the Barkhausen-Kurz connection in oscillation-generating power. For given inter-kathode distance, the oscillations generated under specified circuit conditions have frequency ranges which depend on the relative values of the kathode-grid and grid-grid distances. The lowest frequencies are obtained when the latter distance is large compared with the former; another range of higher frequency appears when the kathode-grid distance is greater than the grid-grid distance. If this disparity is large, a further range of oscillations of still higher frequency can be obtained.

An examination of the results shows definite relationships between applied voltage, electrode distances and wave-length of the ultra-high-frequency oscillations generated. It is possible that the recorded results could be explained by a modification of the Benham-Müller theory which predicts ranges of negative conductance for a diode to which ultra-high-frequency potentials are applied.

§1. INTRODUCTION

The results which are recorded here have been obtained with the demountable tube shown to the Physical Society on 26 November 1937 (Leyshon, 1938). It had previously been found possible to generate ultra-high-frequency oscillations with a sealed tube containing an electrode system of two grids at the same positive D.C. potential with kathodes symmetrically placed externally to them (Leyshon, 1936). The demountable tube was set up in order to examine in greater detail the oscillation-generating properties of such an arrangement; it was also hoped that in the course of the investigation some light might be thrown upon the mode of maintenance of these oscillations.

The apparatus used in the experiments to be described has the advantages of flexibility and ease of alteration of design which all such demountable tubes possess. The construction is as simple as possible; particular care has been taken to avoid effects due to unwanted resonances, and to make the Lecher-wire system which is attached to the grids practically continuous up to the grids themselves.

§2. APPARATUS

Electrode system and tube construction

For convenience, some of the figures given in the account of the demonstration already referred to (Leyshon, 1938) are reproduced in figure 1, which shows the structure of the tube and its parts.

In fitting up the tube the procedure was as follows:—The electrode leads were screwed on to the under side of the Mycalex disc through which passed six screws, held in position by nuts (figure 1 *a*). The filament frameworks (figure 1 *c*) were fitted with tungsten wires, and the grid frameworks (figure 1 *b*) with copper gauze. They were then attached to the electrode leads and adjusted in position so that the electrodes were symmetrically disposed. Careful measurements were

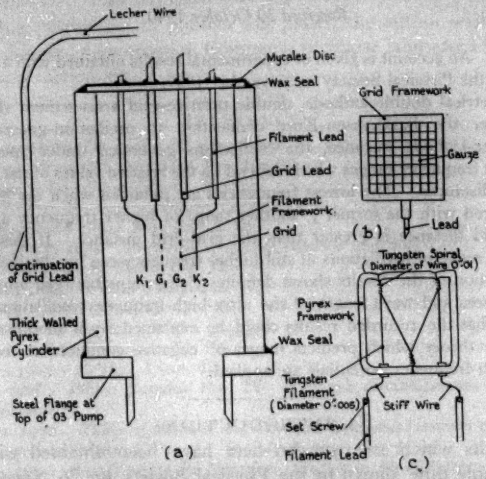


Figure 1. Tube and electrode structure.

made of the electrode distances. The Mycalex disc having been placed in position, the backing pump started, and the Lecher wires connected, all joints were sealed with Apiezon black wax. The tube was then evacuated by a two-stage oil diffusion pump. The electrode system and leads were freed from gas by radiation from the hot filaments, and by electron and positive-ion bombardment, the latter process being facilitated (when the vacuum was high enough) by the effect of ultra-high-frequency oscillations generated by the tube and associated circuit.

Lecher wires

In the final form of the apparatus the Lecher wires used were copper tubes of the same diameter (4 mm.) and at the same central distance apart (2 cm.) as those used for the electrode leads inside the valve. It was possible, with the

rather thick electrode leads, to dispense with any special cooling device for the wax seals. These thick leads * and the desire for continuity involved the use of external Lecher wires of a somewhat larger ratio of diameter to distance apart than would be considered ideal for ease of oscillation. However, the advantage of dispensing with cooling metal rings and their possible resonances outweighed this drawback.

Electrical circuit

In the electrical circuit used in obtaining the tube characteristics, and for production of oscillations, use was made of the 220 v. D.C. mains to provide both filament current and high tension, or part of the latter. The possibility of high-frequency variations of kathode potential due to standing waves on the filament connecting-wires was minimized by connecting the negative ends of the filaments to each other directly, and the positive ends through a 0·003 mfd. condenser. These precautions were taken since previous experiments with a T.M.C. valve had shown that strong oscillations could be set up in Lecher wires in the filament circuit, battery connections being made to the grid and anode by short straight wires.

§ 3. EXPERIMENTAL PROCEDURE AND RESULTS

For convenience of reference, the dimensions of the various electrode systems investigated are given in table 1.

When the tube had been fitted up and the parts freed from occluded gas, the characteristic {log grid voltage, log grid current} curves for the non-oscillating condition were drawn. These were all straight lines with a slope of 3/2 for space-charge-limited currents.

With any given set of internal conditions, the bridge on the Lecher wires was pushed in until the onset of oscillation was made evident by a change in grid current (Δi_g), or by the deflection (D) of the galvanometer in a Lecher-wire-crystal wave-meter, or in an aperiodic crystal circuit coupled to the tube circuit.

Oscillations having been detected, curves of Lecher-wire length (l) versus Δi_g and/or D were obtained for various circuit conditions. The usual procedure was to set the filament current (or currents) at a particular value and then to find these oscillation curves for a series of voltages, the whole process being repeated for different values of the filament current. Some typical results are shown in figure 2. The curves have the same form in whatever way the oscillations were generated, except in the cases of 7 and 14, for which the (l, D) graphs approximate more closely to normal resonance curves.

Wave-length determinations. The wave-length of the generated oscillations was measured either by a separate wave-meter coupled to the circuit or by means of the Lecher wires attached to the grids.

* It is perhaps advisable that these leads and the grids should be silvered, to avoid filament "poisoning" due to sputtered copper.

The wave-meters used were either of the absorption type (Lecher wire or closed circuit (Moore, 1934)) or Lecher wires with crystal detector and galvanometer (Chapman, 1932). The reproducibility of the wave-length determinations

Table 1

System	Distance between kathodes K_1K_2	Distance between grids G_1G_2	Area of grid	Size of grid mesh	Method of use	Notes
	cm.	cm.	sq. cm.			
1	3·0	1·8	4	Close. 50 % space	One or both filaments emitting Grids at same or different potentials	First experimental tube. Not demountable
2	3·5	2·5	3·1*	,,	One filament emitting. Outer grid at zero or negative potential	Demountable
3	2·8	2·1	,,	,,	As for 2	,,
4	2·8	0·9	,,	,,	No oscillations obtained	,,
5	3·85	2·05	,,	,,	As for 2 and 3	,,
6	4·0	2·1	,,	,,	,,	,,
7	3·6	2·2	4·0†	Open. 60–70% space	Both grids at same positive potential	,,
8	3·45	1·1	,,	,,	As for 7	,,
9	2·9	1·1	,,	,,	As for 1	,,
10	2·5	1·1	,,	,,	As for 7	,,
11	2·85	1·45	,,	,,	,,	,,
12	3·25	1·45	,,	,,	,,	,,
13	3·75	1·45	,,	,,	,,	,,
14	3·5	2·5	,,	,,	,,	,,
15	3·5	1·5	,,	,,	,,	,,
16	4·5	1·1	,,	,,	,,	,,
17	2·48	0·6	,,	,,	,,	,,

* Total area including frame = 5·1 sq. cm.
† ,, ,, ,, ,, = 6·25 ,,

on different occasions is shown in figure 3. The straight-line relationship between Lecher-wire length and wave-length was used to estimate wave-lengths when they were not determined by direct measurement.

Relation between wave-length and amplitude of generated oscillations and circuit conditions

Although the wave-lengths of the oscillations generated in the manner described were determined almost entirely by the setting of the Lecher-wire bridge,

yet, as many other experimenters have found using commercial triodes in a Barkhausen-Kurz circuit, the range of Lecher-wire length over which oscillations occurred was dependent on the electrode spacing, the filament temperature and the voltages applied to the grids.

In experiments using the demountable tube no oscillations were obtained with both grids at the same D.C. potential until copper gauze of wider mesh than that

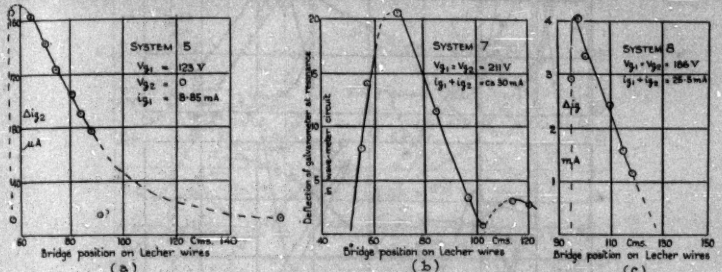


Figure 2. Dependence of oscillation intensity on Lecher-wire length.

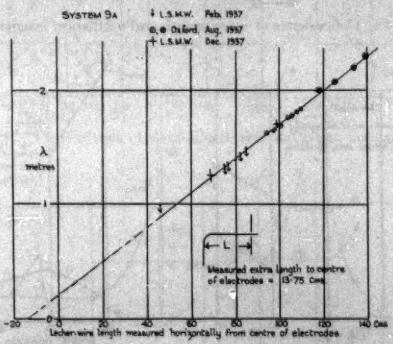


Figure 3. Relation between generated wave-length and Lecher-wire length (9 A).

used in the original tube 1 and systems 2 to 6 was fitted into rather larger grid frameworks. When 7 was set up, however, it was found that oscillations could obtained with the greatest ease.

From graphs like those of figure 2 the values of longest (λ_L), optimum (λ_O) and shortest (λ_S) wave-lengths generated for a particular value of emission and grid voltage were found. System 5 (used with one filament emitting, and the outer grid at filament potential) was studied in detail, and results for this system are given in figure 4.

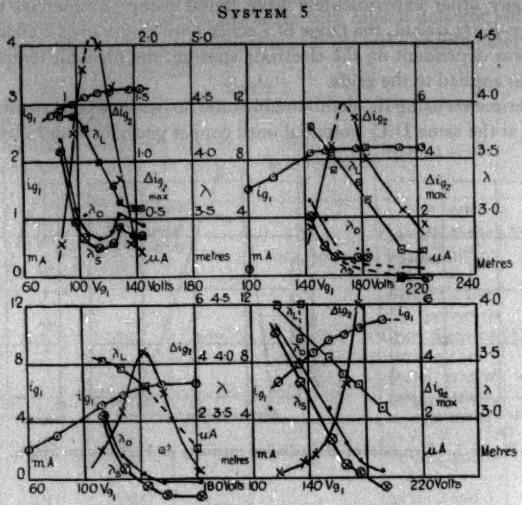


Figure 4. Dependence of wave-length and oscillation intensity on V_g.

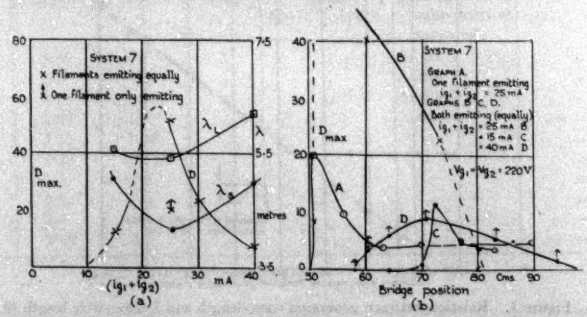


Figure 5.

(*a*) Dependence of wave-length and oscillation intensity on grid current.

(*b*) Dependence of oscillation intensity on Lecher-wire length, for different grid currents.

The effect of grid current (i_g) on the wave-length range and on the oscillation intensity may be shown in an alternative way by plotting (i_g, λ) and (i_g, Δi_g) graphs for a series of constant voltages.* Typical curves are shown in figure 5; from this figure it may be seen that for system 7, used with the same total grid current, the strength of oscillation was considerably reduced when one filament only

* The change Δi_g in grid current is an indication, but not a measure, of the oscillation intensity.

was being used as an emitter. The oscillation amplitude and value of optimum wave-length changed with grid current (the D.C. grid voltage being constant) in the same way as for 5 used as a Barkhausen oscillator.

For systems 7 and 14, for which $G_1G_2 > 2KG$, optimum oscillation occurred when the kathode-grid spaces were carrying emission-limited currents. For

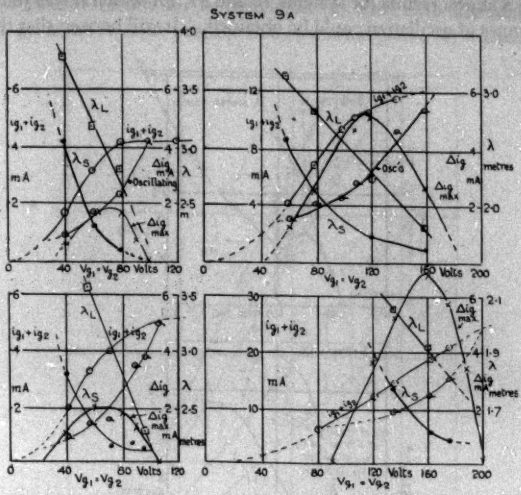


Figure 6. Dependence of wave-length and oscillation intensity on V_{g1}.

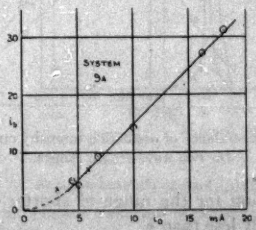


Figure 7. Relation between grid current for optimum oscillation (i_0) and emission-limited current (i_s).

systems 8 to 13 and 15 to 17, in which $G_1G_2 < 2KG$, the (l, D) or $(l, \Delta i_g)$ curves were of the kind shown in figures 2 *a* and 2 *c*; the optimum oscillation occurred for a given voltage when the kathode-grid spaces were carrying space-charge-limited currents. This is shown in figure 6 for system 9 A, which was examined in detail.

In figure 7 the value of the grid current for optimum maintenance of oscillation, i_0, is compared with the emission-limited current i_s for a series of different filament emissions. The values of i_0 and i_s were taken from figure 6. It is clear that for system 9A, i_0 was a linear function of i_s and was nearly $\frac{1}{2}i_s$ for large emissions. In some other cases i_0 had a value relative to i_s greater than this.

Figure 8 shows results for systems 16 and 17, for which it was found that two distinct ranges of oscillation could be obtained. It will be seen that the optimum

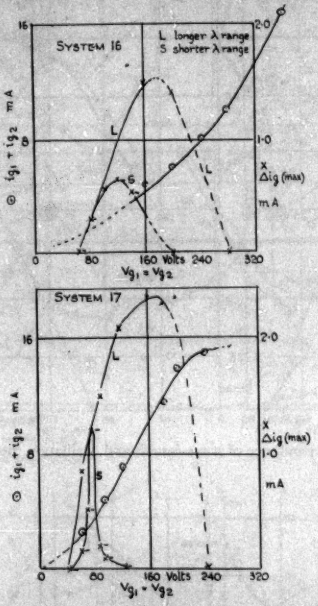


Figure 8. Dependence of oscillation intensity on grid voltage for two wave-length ranges.

voltage for a given emission for the range of shorter wave-length was one for which i_o/i_s was considerably less than $\frac{1}{2}$.

From the $(l, \Delta i_g)$ or (l, D) curves for each setting of grid voltage, results were obtained showing the range of wave-length for which oscillations would occur for this voltage. From a series of such results curves of $\left(\frac{1}{\sqrt{V_g}}, \lambda_L\right)$ and $\left(\frac{1}{\sqrt{V_g}}, \lambda_0 \text{ or } \lambda_S\right)$ were drawn. The maximum range of $\lambda\sqrt{V_g}$ for which oscillations would occur was thus found. In the earlier experiments the range was sometimes deduced from a few observations only. However, as these observations were usually

made under optimum conditions of maintenance for which the greatest range of $\lambda\sqrt{V_g}$ was always found, the complete range for the system under test was probably obtained.

In figure 9 is indicated the degree of accuracy with which this range of $\lambda\sqrt{V_g}$ was found for systems 5 and 9 A, the two systems studied in greatest detail.

Figure 10 shows results for 9 when used (*a*) with both filaments emitting and the two grids at the same positive D.C. potential (9 A); (*b*) with one filament

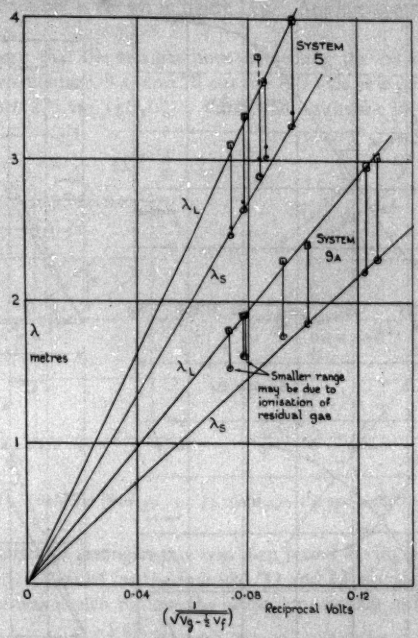


Figure 9. Relation between λ_L (and λ_S) and $\frac{1}{\sqrt{V_g}}$ for optimum conditions of oscillation.

K_1 emitting and the further grid G_2 at the same D.C. potential as the negative end of K_1 (9 B). When only a few observations had been made with arrangement 9 B, it appeared that changing over from (*a*) to (*b*) had practically no effect on the wave-length generated. The final results given in figure 10 *b* show that the range of $\lambda\sqrt{V_g}$ for 9 B is slightly smaller than that for 9 A and the actual values rather less. It should be observed that a condenser bridge had to be used on the Lecher wires in the case of 9 B.

In one particular experiment, the nearer condenser bridge on the Lecher wires was kept fixed in position, while the outer one was adjusted for optimum deflection of the galvanometer in a coupled aperiodic crystal circuit. The emission from one filament being constant, and from the other zero, the voltage on the nearer grid was varied, the outer grid being either (*a*) at the same potential as the nearer or (*b*) at mean filament potential. It was found that the optimum voltage altered little in changing over from (*a*) to (*b*), but the amplitude of the

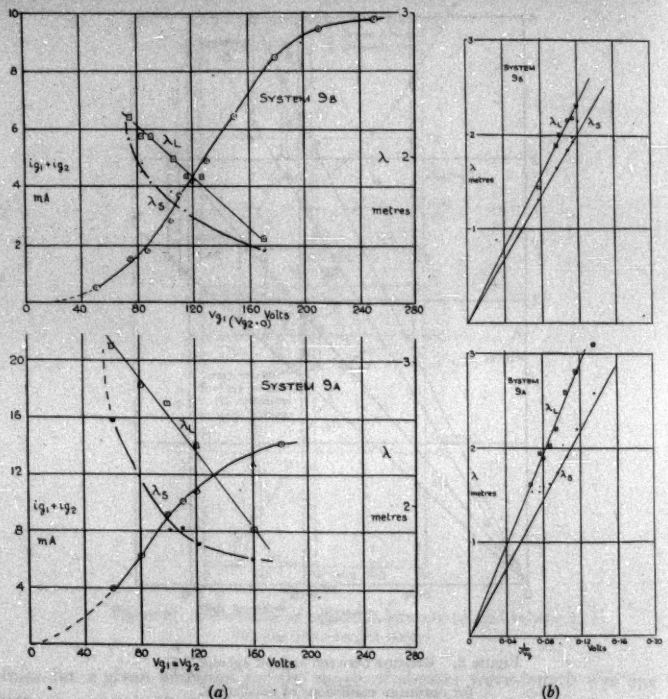


Figure 10. Comparison of systems 9A and 9B: wave-length dependence on V_g

optimum oscillation was diminished. Since with arrangement (*a*) still greater amplitude of oscillation was attained with both filaments emitting, the advantage of the positive double-grid system over the Barkhausen connection is evident.

Finally, having obtained the maximum range of $\lambda\sqrt{V_g}$ for each system, an attempt was made to relate these results to the values of the electrode distances.

For convenience, the results are summarized in table 3 (see page 154), from the first five columns of which it may be seen that for 8 to 13 and 15 to 17 the values of $\lambda\sqrt{V_g}$ are smaller for given electrode distances than for 1, 7, and 14.

In view of the fact that Chipman (1935), in his experiments with a demountable plane-electrode triode, had come to the conclusion that oscillations could be obtained whose period was approximately twice the transit time of an electron from filament to grid, a graph was drawn of $\lambda\sqrt{V_g}$ versus kathode-grid distance KG for 8 to 10. This is shown in figure 11*a*. Arrangements 11 to 17 had not then been set up.

It will be seen that the straight lines connecting the extreme points on the $\lambda\sqrt{V_g}$ range for systems 8, 9 A, and 10 cut the KG axis at a point corresponding to a distance of 0·275 cm. ($\frac{1}{4}G_1G_2$). Later, the results for 16 were added to the graph.

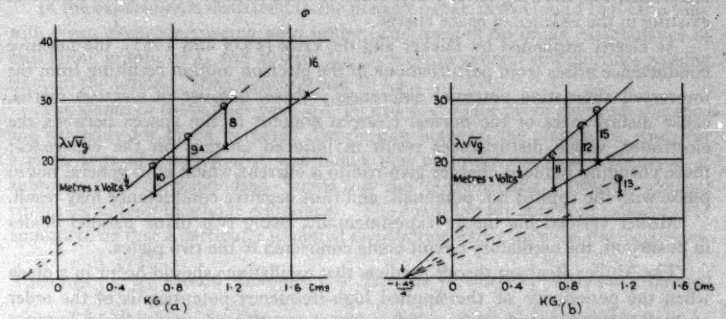


Figure 11. Relation between $\lambda_L\sqrt{V_g}$ (and $\lambda_S\sqrt{V_g}$) and kathode-grid distance.

A further series of arrangements was then tested having a different value of G_1G_2 (1·45 cm.). Results for the first two (11 and 12) are shown in figure 11 *b*. The upper line was drawn to have the same slope as that in figure 11 *a*, and to pass through the point $\left(\frac{-1{\cdot}45}{4}, 0\right)$. It will be seen that the points corresponding to the two upper extreme values of $\lambda\sqrt{V_g}$ lie on this line within the limits of experimental error, and that the line joining the points corresponding to the lower extreme values passes through this point. Results for 15 were added later. Thus for 8 and 9 A to 15, $\lambda_L\sqrt{V}$ and $\lambda_S\sqrt{V_g}$ are each approximately proportional to $KG+1/4\ G_1G_2$.

The first ranges of oscillation for 16 and 17 evidently corresponded to those of 8 and 9A to 15, the second probably to that of 13, on which only a few observations were made.

§4. DISCUSSION OF RESULTS

Benham (1931 and 1938), Müller (1933 and 1934) and others (e.g. Backer and de Vries, 1934 and 1935) have proved theoretically that a diode should show ranges of negative conductance towards a small alternating potential difference impressed upon the D.C. potentials applied to the electrodes. The first range of negative conductance should, according to Benham, occur when the time of transit t of the electrons between the kathode and plate of the diode lies between about 0·97 and 1·31 times the periodic time T of the applied high-frequency potentials, the maximum negative conductance occurring in the neighbourhood of the highest frequency maintainable by the system. Negative conductance may arise whether space charge is, or is not, present to an extent sufficient to affect appreciably the potential distribution between kathode and plate. The time of transit to be considered is that calculated in absence of ultra-high-frequency potentials on the electrodes, but allowing for any departure of the potential distribution from that existing in the absence of space charge.

As clearly explained by Bakker and de Vries (1934 and 1935), the negative conductance arises from perturbations in the electron motion resulting from the impressed alternating potential differences; these, because of electron inertia, cause disturbances of the normal electron content of the spaces between the electrodes, which disturbances result in induced charges on the electrodes; these changing induced charges give rise to a current which is in general not in phase with the applied h.f. potentials, and thus negative conductance may result.

Müller verified his theory experimentally, using two plane parallel diodes in push-pull, the oscillatory circuit being connected to the two plates.

The Müller-Benham theory predicts that oscillations should occur in a diode when the periodicity of the applied high-frequency potentials is of the order of the time of transit of an electron between these electrodes.

Thus, according to Benham's theory, as well as to most other theories which have been put forward to explain Barkhausen-Kurz and similar oscillations, the calculated values of the periodicities of the applied potentials for which the tubes in question show negative conductance bear a definite relationship to the time of transit of an electron between electrodes in the absence of oscillations.

In general, $t = kT$, where k is a factor of proportionality which has a definite range of values for each set of internal conditions.

Theoretically, the time of transit of an electron between kathode and plate lies between the values (i) $\sqrt{\frac{m}{2e}} \cdot \frac{1}{\sqrt{V}} \cdot 2d$ and (ii) $\sqrt{\frac{m}{2e}} \cdot \frac{1}{\sqrt{V}} \cdot 3d$. (i) holds when the influence of space charge on potential distribution is negligible, (ii) holds for the distribution corresponding to space-charge saturation, if the effect of initial velocities is neglected.

If λ = wave-length *in vacuo* of electromagnetic waves of periodicity T,

$\lambda = cT = \frac{ct}{k}$, where c = velocity of light or $\lambda = \frac{c}{k}\sqrt{\frac{m}{2e}} \cdot \frac{1}{\sqrt{V_d}} \cdot \frac{d}{1-a/2}$,* where $\frac{V_d}{x^a}$ = constant represents the distribution of potential between kathode and plate, distant d from each other, x being the distance from the kathode of the point at potential V and

$$\frac{\lambda\sqrt{V_d}}{d} = c/k\sqrt{\frac{m}{2e}} \cdot \frac{1}{1-a/2}.$$

If d and a are constant, then the values of $\lambda\sqrt{V}$ should lie within definite ranges determined by k.

Relation between $\lambda_L\sqrt{V_g}$ and $\lambda_S\sqrt{V_g}$ for given d and given space-charge conditions

In the experiments described in the present paper, it was found that for a tube with given electrode spacing used under conditions of optimum maintenance of oscillation, graphs of λ_L and λ_S with $1/\sqrt{V_g}$ were in general straight lines passing through the origin. (See figures 9 and 10 *b*.)

The slopes of these lines gave $\lambda_L\sqrt{V_g}$ and $\lambda_S\sqrt{V_g}$ for given electrode distance and probably for given potential distribution between the electrodes, since for a given type of oscillation it was found that optimum oscillation occurred at approximately corresponding points on the characteristic {current, voltage} curves for different kathode emissions. The condition for optimum maintenance appears to be that a particular potential distribution exists between the electrodes. The ratio $\frac{\lambda\sqrt{V_g}}{\lambda_S\sqrt{V_g}}$ should therefore give the ratio $\frac{k_2}{k_1}$, where $k_1 \to k_2$ is the predicted range of k. Benham gives the range of k as 0·97 to 1·31, and thus $k_2/k_1 = 1{\cdot}35$.

Experimental values for the different arrangements are given in table 2.

Table 2

System $\lambda_L\sqrt{V_g}/\lambda_S\sqrt{V_g}$		System $\lambda_L\sqrt{V_g}/\lambda_S\sqrt{V_g}$		System $\lambda_L\sqrt{V_g}/\lambda_S\sqrt{V_g}$	
1	1·42	8	1·32	13	1·17
7	1·45	9 *a*	1·33	16 (ii)	1·42
14	1·49	9 *b*	1·21	17 (ii)	1·29
2	1·24	10	1·36		
3	1·31	11	1·37		
5	1·27	12	1·42		
6	1·50	15	1·46		
		16 (i)	1·44	Mean	1·37
		17 (ii)	1·51		

* $t = \int_0^d \frac{dx}{v}$, where v = velocity of electron at distance x from kathode. Since $\frac{1}{2}mv^2 = eV_x$, $v = \sqrt{\frac{2eV_x}{m}}$; also $\frac{V_x}{x^a} = \frac{V_d}{d^a}$. Thus $t = \sqrt{\frac{m}{2e}} \cdot \frac{d^{a/2}}{\sqrt{V_d}} \int_0^d x^{-a/2}dx = \sqrt{\frac{m}{2e}} \frac{1}{\sqrt{V_d}} \cdot \frac{d}{1-a/2}$.

It is seen that the mean value differs little from that predicted. The larger values of the ratio correspond in general to those cases in which the shortest waves were of the " non-restoring " type.

Relation between $\lambda\sqrt{V_g}$ and effective electrode distance

If we follow Benham and assume that the direct time of transit t is 0·97 to 1·31 times the periodicity T of the first range of generated oscillations, we can find from the experimental results the distance d between the kathode and plate of the temperature-limited diode equivalent to the system actually used. The distance between kathode and plate of the corresponding space-charge-limited diode would be two-thirds of this.

Putting λ in metres, V in volts, c in metres per second, $\sqrt{\frac{m}{2e}}$ in electromagnetic units, and $a = 1$ in the formula previously given:—

$$d\ (\text{cm.}) = \lambda\sqrt{V} \times k/10{\cdot}1.$$

Thus $$d(\lambda_L) = \lambda_L\sqrt{V_g} \times 0{\cdot}97/10{\cdot}1$$

and $$d(\lambda_S) = \lambda_S\sqrt{V_g}1 \times {\cdot}31/10{\cdot}1.$$

In table 3 are listed the values (to the nearest mm.) found in this way for the electrode distances of the temperature-limited diodes equivalent to the various systems tested.

Table 3

Arrangement	KG (cm.)	G_1G_2 (cm.)	$\lambda_L\sqrt{V_g}$	$\lambda_S\sqrt{V_g}$	$d(\lambda_L)$ (cm.)	$d(\lambda_S)$ (cm.)
1	0·6	1·8	64	45	6·1	5·85
7	0·7	2·2	80	55	7·7	7·2
14	0·5	2·5	90	60·5	8·6	7·9
2	0·5	2·5	54	43·5	5·2	5·7
3	0·35	2·1	52·5	40	5·0	5·2
4	0·95	0·9	no oscillations detected			
5	0·925	2·05	44·5	35	4·3	4·6
6	0·95	2·1	54	36	5·2	4·7
9 B	0·9	1·1	20	16·5	1·9	2·1
8	1·15	1·1	29	22	2·8	2·9
9 A	0·9	1·1	24	18	2·3	2·3
10	0·65	1·1	19	14	1·8	1·8
11	0·7	1·45	21	15	2·0	2·0
12	0·9	1·45	26	18	2·5	2·3
15	1·0	1·45	28·5	19·5	2·7	2·5
16 (i)	1·7	1·1	44·5	31	4·3	4·0
17 (i)	0·94	0·6	23	15	2·2	2·0
13	1·15	1·45	17	14	1·6	1·8
16 (ii)	1·7	1·1	25	18	2·4	2·3
17 (ii)	0·94	0·6	13	10	1·25	1·3

In table 4 the values of d shown in table 3 are compared with the electrode distances of the various systems studied. It will be seen from this table that for 1, 7, and 14

$$d = 2(2KG + G_1G_2) \text{ or } 3{\cdot}5\ G_1G_2.$$

For 8 to 17, omitting 13, 16 (ii), 17 (ii),

$$d = 2KG + \tfrac{1}{2}G_1G_2 \text{ or } 2{\cdot}08\,(KG + 0{\cdot}24),$$

while for 13, 16 (ii), 17 (ii)

$$d = 1{\cdot}42\ KG.$$

It should perhaps be re-stated here that the optimum conditions of oscillation maintenance for 1, 7, and 14 occurred when the grid current was emission-limited, for 8 to 17 when this current was space-charge-limited, and for 16 (ii) and 17 (ii) when this current was greatly limited by space charge.

Table 4. Values of d (cm.)

System	$d(\lambda_S)$	$2(2KG+G_1G_2)$	$d(\lambda_L)$	$3{\cdot}5\ G_1G_2$
1	5·85	6·0	6·1	6·3
7	7·2	7·2	7·7	7·7
14	7·9	7·0	8·6	8·75
	$d(\lambda_S)$	$2{\cdot}08\ (KG+0{\cdot}24)$	$d(\lambda_L)$	$2KG+\frac{1}{2}G_1G_2$
9 A	2·3	2·4	2·3	2·35
8	2·9	2·9	2·8	2·85
10	1·8	1·85	1·8	1·85
11	2·0	2·0	2·0	2·1
12	2·3	2·4	2·5	2·5
15	2·5	2·6	2·7	2·7
16 (i)	4·0	4·0	4·3	4·0
17 (i)	2·0	2·45	2·2	2·2
	$d(\lambda_S)$	$1{\cdot}42\ KG$	$d(\lambda_L)$	
13	1·8	1·6	1·6	
16 (ii)	2·3	2·4	2·4	
17 (ii)	1·3	1·3	1·25	

One fact which emerged particularly clearly in the course of this investigation was that arrangements with corresponding electrode distances behave similarly.

1 and 7, for which $G_1G_2/KG = 3$ and 3·1 respectively, gave the same range of $\lambda\sqrt{V_0}/d'$, d' representing corresponding electrode distances in the two cases. 1 had fine-meshed grids, 7 open-meshed grids; the Lecher-wire system used with 1 was different from that used for 7 and all subsequent arrangements.

This suggests that a single electrode space may not be the seat of the oscillations, but that the ease of generation of oscillation and the wave-length depend on the relative electrode distances.

Since, also, the ratio $\dfrac{\lambda_L\sqrt{V_0}}{\lambda_S\sqrt{V_0}}$ was of the same order for 1, 7, and 14 and for 2 to 6 as for the other systems, it seems likely that the same type of mechanism of maintenance was operative in all the systems studied. The difference in the

relation between d and electrode distances for the different systems may perhaps be attributed to effects of different space-charge distribution in the electrode interspaces, and also possibly to different values of effective grid-transparency, on electron transit time; further, the electron paths may not be straight lines perpendicular to the planes of the electrodes.

In calculating transit times, the effective voltages at a kathode due to both grids would have to be taken into account for 8 to 17; for 1, 7, 14 and 2 to 6 the effect of the outer electrode would probably be negligible—in the first case because the G_1G_2 space was saturated with space charge, in the second because the outer grid was at zero potential. The effect of initial velocities, and of the non-planar character of the kathodes, would also have to be considered in a complete theoretical discussion of such transit times.

That the oscillation-maintenance could possibly be accounted for by a modification of the Benham-Müller theory is suggested by the definite relation found experimentally between inter-electrode distance and plate-kathode distance of the equivalent diode calculated from this theory. On the other hand, the values for the kathode-plate distance for the equivalent diodes calculated from Benham's theory are, particularly for 1, 7, and 14, much larger than would be expected for direct electron transits between the electrodes.

It is interesting to note that Benham's values for pt (where $p=2\pi/T$) for the first and second optima of negative conductance for a diode are 7·6 and 13·85 respectively, and thus, for identical conditions of operation, the optimum wavelengths generated should have the ratio $13{\cdot}85/7{\cdot}6=1{\cdot}82$. The ratio found experimentally for 16 was actually 1·81.

§ 5. CONCLUSION AND ACKNOWLEDGEMENTS

It has not been possible in the present circumstances to carry this experimental investigation further. I should like to express my thanks to Professor J. S. Townsend, F.R.S., for his kindness in allowing me to work during summer vacations in the Electrical Laboratory at Oxford; to the Council of the London (R. F. H.) School of Medicine for Women (in the Physics Laboratory of which the greater part of the work was done at intervals during 1937–9) for a grant from the Augustus and Alice Waller Memorial Research Fund; and to Dr. W. H. Eccles, F.R.S., who suggested the construction of the first tube examined and who has taken a continued interest in the experiments.

REFERENCES

Bakker, C. J., and de Vries, G., 1934. *Physica*, **1**, 1045 ; 1935. *Ibid.* **2**, 683.
Benham, W. E., 1931. *Phil. Mag.* **9**, 457 ; 1938. *Proc. Inst. Radio Engrs*, **26**, 1093.
Chapman, F. W., 1932. *Wireless Engr.* **8**, 500.
Chipman, R. A., 1935. *Proc. Phys. Soc.* **47**, 263, 1042.
Leyshon, W. A., 1936. *Proc. Phys. Soc.* **48**, 469.
Leyshon, W. A., 1938. *Proc. Phys. Soc.* **50**, 146.
Moore, W. H., 1934. *J. Franklin Inst.* **217**, 347.
Muller, J., 1933. *Hochfrequenztechnik u. Elektroakust.* **41**, 156 ; 1934. *Ibid.* **43**, 195.

NOTE ON THE BEST FOCUS IN THE PRESENCE OF SPHERICAL ABERRATION

By WANG TA-HANG,
Imperial College of Science and Technology

Communicated by L. C. Martin. Received 14 *November* 1940

ABSTRACT. The location of the best axial focus of an optical system in the presence of spherical aberrations of various orders is considered upon the criterion of maximum axial light intensity. Adhering to Rayleigh's quarter-wave-length limit, it is shown by approximate formulation that for axial aberrations sensible up to the primary, secondary and tertiary order, the allowable tolerance for such parts of the aberrations when stated in terms of optical path-difference can reach λ, 4λ, and 15λ respectively, provided that the foci of various zones are suitably controlled in design.

§1. INTRODUCTION

THE present paper deals with the location of the best focus of an axial pencil of rays in the image space when residual spherical aberrations of various orders are present. The treatment is based on the criterion that the best location is the point where the axial light intensity is a maximum. Cases are considered when spherical aberrations of various orders are under control in practical designing work. Rayleigh's quarter-wave-length limit is adhered to because an optical instrument would fall very little short of its utmost theoretical perfection if all the light arrived at the focus with differences of optical path not exceeding that limit. This has been proved both mathematically (Conrady, 1919; Buxton, 1921; Martin, 1926) and experimentally (Conrady, 1926). It also simplifies the mathematical treatment given in the present paper. Results obtained are compared with those obtained by other writers and published elsewhere. Methods applied to practical design are also discussed, with special emphasis on the case of tertiary aberration.

§2. THEORY

Let B in figure 1 represent a point on the optical axis near the final focus in the image space of an axially symmetrical optical system. AP is a spherical surface concentric with B. We can consider this as a reference surface, with which the wave front (originating from an axial luminous object point and distorted on passing through this optical system) can be compared. In accordance with the notation used by Professor A. E. Conrady, the product of the radial

distance from the reference surface to the wave front and the refractive index in the space concerned is termed the optical path difference, written symbolically as OPD. If y is the intersection height of the ray with this reference surface (or a quantity proportional to the numerical aperture of the ray) and y_0 that at the extreme margin, then OPD can be put in the form of a power series:

$$\mathrm{OPD}_B = a_0 + a_1\left(\frac{y}{y_0}\right)^2 + a_2\left(\frac{y}{y_0}\right)^4 + a_3\left(\frac{y}{y_0}\right)^6 + \ldots, \qquad \ldots\ldots(1)$$

where the suffix B means a value taken with respect to a reference surface with B as centre. Professor Conrady's treatment disregarded the constant term a_0 in the above expression. In that case the wave front has its vertex in contact with the reference surface. We will, however, retain this constant term for a reason to be given below, but at present we restrict it to the magnitude of the order of a few wave-lengths of light, so that it becomes comparable in magnitude with the other terms in the same expression.

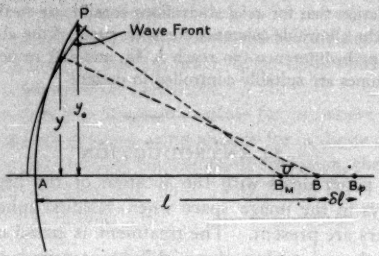


Figure 1.

Imagine the point B to lie in the vicinity of the best focus. It follows that the various terms in expression (1) will have the following meanings :

a_0 term. As pointed out above, this term is related to the proper reference surface chosen. The choice of its value is arbitrary and there is no effect on the quality of the image formation, but we shall see in the following sections how a proper choice of this term leads to simplification of the mathematical treatment, where a_0 is set at a value representing the " mean phase " of the wave front.

$a_1(y/y_0)^2$ term. This term is related to the position of the point B and is the only sensible term representing the focal shift, provided that B is very near to the focus and the angle of the cone of rays is not too big. To see this point, note that at the paraxial focus, the OPD expression starts from the term in y^4 (disregarding the constant term a_0). It is not difficult to show, using Fermat's principle and simple trigonometrical deductions, that that part of OPD caused by a focal shift δl from the paraxial focus is given by

$$\mathrm{OPD(focal\ shift)} = 2N\delta l \sin^2\frac{U}{2},$$

provided that $(\delta l)^2$ is negligible, where U is the angle subtended by the height of y from B. Expanding the above expression in powers of $\sin U$ and replacing $\sin U$ by y/l, we obtain

$$\text{OPD (focal shift)} = \tfrac{1}{2}N\delta l\,(\sin^2 U + \tfrac{1}{4}\sin^4 U + \tfrac{1}{8}\sin^6 U + \dots)$$
$$= \tfrac{1}{2}N\delta l\left[\left(\frac{y}{l}\right)^2 + \tfrac{1}{4}\left(\frac{y}{l}\right)^4 + \dots\right].$$

In practice, for an optical instrument with Rayleigh-limit definition, we can very nearly always find one space, either the object space or the image space or spaces between the lenses, in which the marginal U-value does not exceed 1/8 radian. For instance, for a microscope objective, the marginal U-value at the image side usually does not exceed 1/30 radian; for a telescope of considerable light-gathering power the relative aperture is of the order of f/5, corresponding to a marginal U of 1/10 radian. Let us refer hereafter to such a space. Consider the case having the biggest possible value of U in the above expression, say $U = 1/8$. We see that the second term is only 1/250 of the first term, and the third term is still smaller. Hence we have good reason to say that the first term is the only sensible part in the OPD expression affected by focal shift, provided the above-mentioned restriction has been adhered to. Thus

$$a_1\left(\frac{y}{y_0}\right)^2 = \frac{N}{2}\delta l\frac{y^2}{l^2} \quad \text{or} \quad a_1 = \delta l\frac{N}{2}\frac{y_0^2}{l^2}. \qquad \dots\dots(2)$$

Note that a_1 is proportional to the focal shift δl.

One may also notice in the later sections, where spherical aberrations of various orders are dealt with, that it is justifiable to neglect the higher-order terms for the focal shift.

$a_2(y/y_0)^4$ *term.* This term represents the primary spherical aberration.

$a_3(y/y_0)^6$, $a_4(y/y_0)^8$, *etc.* These are terms representing secondary, tertiary, and other higher-order aberrations respectively.

Consider the light intensity at the point B resulting from disturbances derived from each zone at the reference surface, extending from the optical axis to the full aperture. For practical purposes we may assume that the amplitude of the disturbances over the reference surface is constant for all zones, but the phase differs from zone to zone by the amount $\dfrac{2\pi\text{OPD}_\text{B}}{N\lambda} = \dfrac{2\pi\text{OPD}_\text{B}}{\lambda_0}$, λ being the wavelength in the space concerned and λ_0 that in air. Again, confining ourselves to small apertures as before, by Huygens' principle we can express the disturbances at B at any instant t by the complex expression

$$S_\text{B} = \frac{2\pi A}{\lambda l}\int_0^{y_0} e^{i2\pi\left[\left(\frac{t}{T}-\frac{l}{\lambda}\right)+\frac{\text{OPD}_\text{B}}{\lambda_0}\right]} y\,dy$$
$$= \frac{2\pi A}{\lambda l}e^{i2\pi\left(\frac{t}{T}-\frac{l}{\lambda}\right)}\int_0^{y_0} e^{i2\pi\frac{\text{OPD}_\text{B}}{\lambda_0}} y\,dy.$$

Put $(y/y_0)^2 = z$. Then

$$\left.\begin{aligned} S_B &= \frac{\pi A}{\lambda l} e^{i2\pi\left(\frac{t}{T}-\frac{l}{\lambda}\right)} \int_0^1 e^{i2\pi \frac{\mathrm{OPD}_B}{\lambda_0}} dz \\ &\equiv S_0(\eta + i\xi), \end{aligned}\right\} \quad \ldots\ldots(3)$$

where $S_0 = \frac{\pi A}{\lambda l} e^{i2\pi\left(\frac{t}{T}-\frac{l}{\lambda}\right)}$ is the disturbance at B, assuming perfect axial image-formation, when the OPD for all zones vanishes over the whole aperture, and

$$\eta = \int_0^1 \cos 2\pi \frac{\mathrm{OPD}_B}{\lambda_0} dz, \quad \xi = \int_0^1 \sin 2\pi \frac{\mathrm{OPD}_B}{\lambda_0} dz. \quad \ldots\ldots(4)$$

Thus the amplitude given by the present system as compared with the ideal one bears the ratio

$$\frac{S_B}{S_0} = \eta + i\xi,$$

and the corresponding ratio in intensity measure is given by

$$R^2 \equiv \left|\frac{S_B}{S_0}\right|^2 = \eta^2 + \xi^2. \quad \ldots\ldots(5)$$

Equations (5), (4), and (1) present R as a function of the coefficients a_1, a_2, a_3, etc. In practical designing work, such as microscope-objective design, where the best axial focus is aimed at, once the general arrangement is assigned, the main object to be achieved consists in adjusting the optical system in all admissible ways so that the controllable aberrations of various orders tend to make the best compromise against the residual uncontrollable ones, i.e. R must be a maximum. Thus, regarding a_j, a_k, etc., as parameters, the above criterion requires that

$$\frac{\partial R}{\partial a_j} = 0, \quad \frac{\partial R}{\partial a_k} = 0, \text{ etc.} \quad \ldots\ldots(6)$$

The straightforward computation of equation (4) in connection with (6) is rather elaborate. For practical purposes let us introduce certain approximations in the following manner :—

Expand sin and cos $\frac{\mathrm{OPD}_B}{\lambda_0}$ and put into equations (4):

$$\eta = \int_0^1 \left[1 - \frac{1}{2!}\left(\frac{\mathrm{OPD}_B}{\lambda_0}\right)^2 + \frac{1}{4!}\left(\frac{\mathrm{OPD}_B}{\lambda_0}\right)^4 - \ldots\right] dz,$$

$$\xi = \int_0^1 \left[\frac{\mathrm{OPD}_B}{\lambda_0} - \frac{1}{3!}\left(\frac{\mathrm{OPD}_B}{\lambda_0}\right)^3 + \ldots\right] dz.$$

Remembering that in equation (1) the constant a_0 is under our arbitrary choice, set a_0 such that

$$\int_0^1 \mathrm{OPD}_B\, dz = 0. \quad \ldots\ldots(7)$$

Let the OPD expression bearing the condition (7) be distinguished by the symbol,*

Then we have

$$\left.\begin{aligned}\eta &= \int_0^1 \left[1 - \frac{1}{2!}\left(\frac{\mathrm{OPD}_B^*}{\lambda_0}\right)^2 + \frac{1}{4!}\left(\frac{\mathrm{OPD}_B^*}{\lambda_0}\right)^4 + \cdots\right] dz, \\ \xi &= -\int_0^1 \left[\frac{1}{3!}\left(\frac{\mathrm{OPD}_B^*}{\lambda_0}\right)^3 - \cdots\right] dz.\end{aligned}\right\} \quad \ldots\ldots(8)$$

What is the physical significance of condition (7)? Referring to figure 2, the curve O′ABC represents the OPD at various zones extending from $z=0$ to $z=1$. Condition (7) means a choice of the axis OZ such that the shaded areas between the curve and the axis on both sides of the latter are equal. Thus OZ is at the " mean " phase of the curve. It will not be far from midway between the extreme values of OPD_B^*. In this diagram these extreme values are at the points O′ and C.

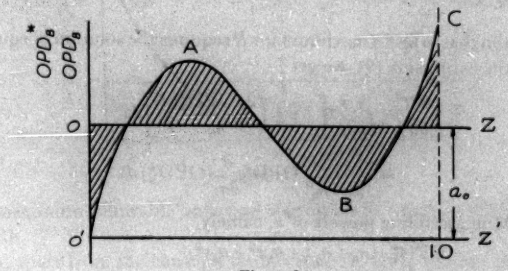


Figure 2.

Having specified the starred OPD_B^* value as above, to guard against ambiguity we will hereafter use the plain OPD for those expressions having the constant term $a_0=0$. Thus OPD_B refers to the axis O′Z′ where $\mathrm{OPD}_B=0$ at $z=0$.

If we adhere to Lord Rayleigh's quarter-wave-length limit as the tolerance for the extreme difference of OPD_B^*, based on the above reasoning, the absolute maximum value of OPD_B^* will not be far from $\frac{1}{8}\lambda_0$ when the full tolerance is allowed, whilst a large proportion of the zones will still be covered by values below that amount. When $\mathrm{OPD}_B^*=\frac{1}{8}\lambda_0$, then $\frac{2\pi \mathrm{OPD}_B^*}{\lambda_0}=\frac{\pi}{4}$, and the magnitude of $\cos\frac{2\pi(\mathrm{OPD}_B^*)}{\lambda_0}$ differs from $1-\frac{1}{2}\left(\frac{2\pi \mathrm{OPD}_B^*}{\lambda_0}\right)^2$ by not more than 2 per cent. If we retain only the first two terms in the expression for η under the integral in (8), the error introduced into the evaluation of η will seldom be 2 per cent, and the total error will be sensibly less, owing to contributions of zones at smaller OPD_B^* under the integral, where the approximation would be more accurate.

Consider next the value of ξ. Take the extreme case by putting $\frac{\mathrm{OPD}_B^*}{\lambda_0}=\frac{\pi}{4}$ throughout. ξ will have the value 1/10. Hence $|\xi|<\frac{1}{10}$, $\xi^2<\frac{1}{100}$. The error

in the evaluation of $R^2=\eta^2+\xi^2$ by neglecting ξ^2 would hardly exceed 1 per cent.

Consequently, the error due to the replacement of R in the full expression by the approximate value

$$R=\int_0^1\left[1-\tfrac{1}{2}\left(\frac{2\pi \mathrm{OPD}_{\mathrm{B}}^*}{\lambda_0}\right)^2\right]dz \qquad \ldots\ldots(9)$$

will be sensibly less than 5 per cent, probably below 1 per cent. To illustrate this by an example, the writer has evaluated R for a proper $\mathrm{OPD}_{\mathrm{B}}$ (expression (11) in the next section), allowing full tolerance, first, by the simplified expression (9), and then by the expression (8), taking into account one more term in the integrals η and ξ than in our simplified expression. The difference between these two values of R is only 0·3 per cent. Thus there is good reason to adopt the simplified expression (9) provided Rayleigh's quarter-wave-length limit is allowed as tolerance.

We have seen that to find a maximum for R requires the solution of equation (6). Putting R in the expression (9), we get

$$\frac{\partial R}{\partial a_j}=\int_0^1\frac{\partial}{\partial a_j}\left[1-\tfrac{1}{2}\left(\frac{2\pi \mathrm{OPD}_{\mathrm{B}}^*}{\lambda_0}\right)^2\right]dz$$

$$=\int_0^1-\frac{2\pi^2}{\lambda_0^2}\mathrm{OPD}_{\mathrm{B}}^*\frac{\partial}{\partial a_j}(\mathrm{OPD}_{\mathrm{B}}^*)\,dz.$$

Writing $\mathrm{OPD}_{\mathrm{B}}^*$ in ascending powers of z, namely,

$$\mathrm{OPD}_{\mathrm{B}}^*=a_0+a_1z+a_2z^2+a_3z^3+\ldots,$$

the equation becomes

$$\int_0^1 z^j(a_0+a_1z+a_2z^2+\ldots)\,dz=0, \qquad \ldots\ldots(10)$$

which involves only simple integration. Remember here that condition (7) has to be involved in the calculation,

$$\int_0^1 \mathrm{OPD}_{\mathrm{B}}^*\,dz=\int_0^1(a_0+a_1z+a_2z^2+\ldots)\,dz=0. \qquad \ldots\ldots(7')$$

This condition corresponds to $\partial R/\partial a_0=0$ by (9).

In the following sections we will consider some cases separately.

§ 3. PRIMARY ABERRATION

For systems where the OPD values are well represented by expressions in powers of y up to the fourth order only, while the higher-order terms are insensible throughout the aperture, we can express

$$\mathrm{OPD}_{\mathrm{B}}^*=a_0+a_1z+a_2z^2.$$

This is the case when only the primary aberration is within the scope of consideration. Assume that the primary aberration cannot be controlled. Here the only controllable parameter is the a_1, corresponding to a shift of focus from the paraxial. We require, according to equations (10) and (7'),

$$\int_0^1 (a_0 + a_1 z + a_2 z^2)\, dz = 0,$$
$$\int_0^1 z(a_0 + a_1 z + a_2 z^2)\, dz = 0.$$

Integrate, and solve for a_0, a_1 in terms of a_2:

$$\left.\begin{aligned} a_1 &= -a_2, \quad a_0 = +\tfrac{1}{6}a_2, \\ \mathrm{OPD}_B^* &= \tfrac{1}{6}a_2 - a_2 z + a_2 z^2. \end{aligned}\right\} \qquad \ldots\ldots(11)$$

The shape of OPD_B^* is illustrated in figure 3. The extreme OPD_B^* difference is $a_2/4$. Using $\lambda_0/4$ as tolerance, we should have $|a_2/4| \leqslant \lambda_0/4$ or $|a_2| \leqslant \lambda_0$, indicating

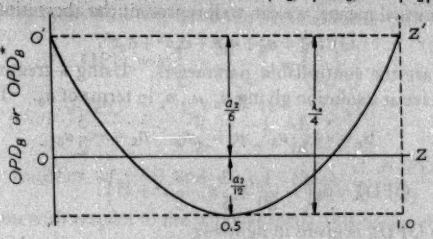


Figure 3.

that the maximum allowable spherical aberration in terms of OPD is a whole wave-length.

In fact, apart from the constant term, expression (11) has just the same form as that obtained by other authors based on the criterion of minimum phase difference (Conrady, 1923; Martin, 1930). Hence the same conclusions can be drawn, namely:

(1) The best physical focus lies midway between the extreme marginal and paraxial foci.

(2) The residual difference of optical paths is only one-quarter of the magnitude attained by them at the extreme marginal or paraxial focus.

(3) The paraxial part of OPD_B^*, namely, $a_2 z^2$, may be allowed to reach a whole wave-length without infringement of the Rayleigh limit, as has been proved above.

Whilst the present author reaches the same conclusion as the others did, the physical meanings are different. Based on the principle of minimum phase-difference, the " best " focus derived can only be admitted with reserve. It is the above formulation that gives the proof.

Let us now calculate the reduction of intensity when the full tolerance is allowed. Using expression (9),

$$R = \int_0^1 \left[1 - \tfrac{1}{2}\left(\frac{2\pi \mathrm{OPD}_B^*}{\lambda_0}\right)^2\right] dz$$
$$= 1 - \tfrac{1}{2}\left(\frac{2\pi}{\lambda_0}\right)^2 a_2^2 \int_0^1 (\tfrac{1}{6} - z + z^2)^2\, dz = 1 - \tfrac{1}{2}\left(\frac{2\pi}{\lambda_0}\right)^2 \frac{a_2^2}{180}.$$

Put $a_2=\lambda_0$. Hence

$$R=1-0{\cdot}11. \qquad \ldots\ldots(12)$$

Thus the central maximum amplitude at the best focus has been reduced from the ideal value by 11 per cent. This accords well with the results obtained from strict numerical work (Conrady, 1919; Buxton, 1921).

§ 4. ZONAL ABERRATION

In an optical system of higher aperture, where the secondary aberration—the $a_3(y/y_0)^6$ term—comes out with sensible magnitude, but higher-order aberrations are still inconsiderable, and the primary aberration is under our control by the designer's usual means, we can well represent the aberration by

$$\mathrm{OPD}_B^*=a_0+a_1z+a_2z^2+a_3z^3.$$

Here, a_1 and a_2 are the controllable parameters. Using a treatment similar to that of § 3, we arrive at a solution giving a_0, a_1, a_2 in terms of a_3. The results are:

$$a_0=-\frac{1}{20}a_3, \quad a_1=\frac{3}{5}a_3, \quad a_2=-\frac{3}{2}a_3,$$

$$\mathrm{OPD}_B^*=a_3\left(-\frac{1}{20}+\frac{3}{5}z-\frac{3}{2}z^2+z^3\right). \qquad \ldots\ldots(13)$$

The shape of the OPD_B^* is given in figure 4.

The value of R can be calculated in a similar manner:

$$R=1-\tfrac{1}{2}a_3{}^2\left(\frac{2\pi}{\lambda_0}\right)^2\int_0^1\left(-\frac{1}{20}+\frac{3}{5}z-\frac{3}{2}z^2+z^3\right)^2dz$$

$$=1-\tfrac{1}{2}a_3{}^2\left(\frac{2\pi}{\lambda_0}\right)^2\frac{1}{2800}. \qquad \ldots\ldots(14)$$

We noticed in the previous section that the reduction of amplitude is 11 per cent

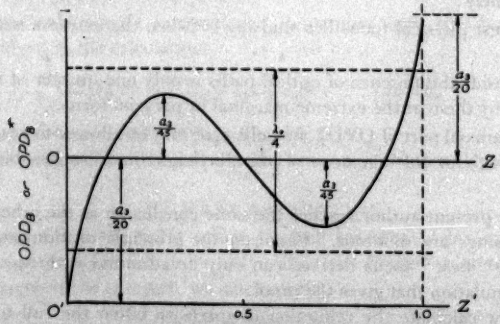


Figure 4.

when the full Rayleigh limit is allowed. In the present case, we may treat the problem in the reverse way. Set the value 11 per cent as the margin for amplitude reduction, below which the image would no longer be considered as perfect.

We can find the allowable values of OPD_B^*. From expression (14), to have $R \geqslant 1{\cdot}0{\cdot}11$ requires $|a_3| \leqslant 4\lambda_0$. In figure 4 the dotted horizontal lines indicate the Rayleigh limit when $|a_3| = 4\lambda_0$, which is the maximum allowable value of $|a_3|$. Note that there are zones near the axis and the extreme margin whose OPD_B^* values exceed $\lambda_0/8$. But as these are only very narrow zones, we need not fear that the use of the simplified expression (9) will cause serious error.

Thus we come to the conclusion:—

(1) With a proper choice of primary spherical aberration and focal shift, the maximum allowable zonal term in OPD_B^* can be as high as 4 wave-lengths.

Take the OPD_p value with respect to the paraxial focus:

$$OPD_p = a_3\left[-\frac{3}{2}\left(\frac{y}{y_0}\right)^4 + \left(\frac{y}{y_0}\right)^6\right],$$

$$\frac{dOPD}{dy} = -\frac{a_3}{y_0}\left[6\left(\frac{y}{y_0}\right)^3 - 6\left(\frac{y}{y_0}\right)^5\right].$$

This derivative vanishes at $y=0$ and $y=y_0$. Since $\frac{1}{N}\frac{dOPD_p}{dy}$ indicates the angular aberration with respect to the paraxial focus, this means that the marginal ray would pass through the paraxial focus. Hence we come to the second conclusion:

(2) In order to get the best axial focus, the extreme marginal ray should be brought to the paraxial focus.

In the expression (13) the existence of the y^2 term means that the best focus is not at the marginal-paraxial focus, but at a focal shift given very nearly by

$$\delta l = -\frac{6}{5}a_3\frac{l^2}{Ny_0^2},$$

by equation (2). Furthermore, put the aberration at various zones in longitudinal measure:

$$B_pB = \frac{l^2}{N}\frac{dOPD_p}{y\,dy} = \frac{2l^2}{Ny_0^2}\frac{dOPD_p}{dz}$$

$$= -a_3\frac{l^2}{Ny_0^2}(6z - 6z^2).$$

B_pB is maximum at $z=0{\cdot}5$ or $y=0{\cdot}707$. Call the most remote zonal focus B_z and we find

$$BB_z = -\frac{3}{2}a_3\frac{l^2}{Ny_0^2} = -\frac{3}{2}\frac{a_3}{N\sin^2 U_0},$$

$$\frac{\delta l}{B_pB_z} = \frac{4}{5},$$

where U_0 is the intersecting angle at full aperture. Thus we come to the third conclusion:

(3) The maximum longitudinal zonal aberration is contributed by the 0.707 zonal ray. The best focus is at 4/5 of the distance from the marginal-paraxial focus to the most remote zonal focus.

Comparing equation (13) with those derived by the same kind of treatment, but based on the principle of minimum phase difference, we find, after a slight transposition to give better accordance with the notations used here, that the expression for OPD_B is given elsewhere (Conrady, 1923; Martin, 1926) by

$$\mathrm{OPD}_B = \frac{9}{16}a_3\left(\frac{y}{y_0}\right)^2 - \frac{3}{2}a_3\left(\frac{y}{y_0}\right)^4 + a_3\left(\frac{y}{y_0}\right)^6.$$

Apart from the constant term, there is a slight difference in the $a_1(y/y_0)^2$ term, which corresponds to a slight shift to the so-called best focus. The other conclusions reached are the same.

§5. TERTIARY ABERRATION

In systems like high-power microscope objectives, the apertures are so big that aberrations of y^8 or even higher orders have to be taken into consideration. The zonal aberration is still controllable in a good design. Thus our controllable

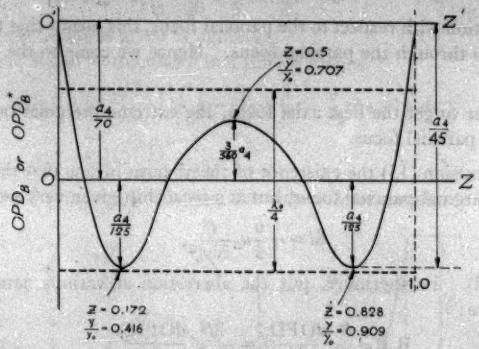


Figure 5.

parameters are a_1, a_2, a_3. At present, as an approximation only, we will assume that only the y^8 term is of sensible magnitude among the higher-order terms. With this restriction, we can obtain a_0, a_1, a_2, a_3 by the same type of treatment as before:

$$a_0 = \frac{a_4}{70}, \quad a_1 = -\frac{2}{7}a_4, \quad a_2 = \frac{9}{7}a_4, \quad a_3 = -2a_4,$$

from which we come to the conclusion:

(1) For best correction, we should arrange the optical system so that the final OPD_B^* comes out as

$$\mathrm{OPD}_B^* = a_4\left[\frac{1}{70} - \frac{2}{7}\left(\frac{y}{y_0}\right)^2 + \frac{9}{7}\left(\frac{y}{y_0}\right)^4 - 2\left(\frac{y}{y_0}\right)^6 + \left(\frac{y}{y_0}\right)^8\right], \quad \ldots\ldots(15)$$

or at the paraxial focus,

$$\mathrm{OPD}_{\mathrm{P}} = a_4\left[\frac{9}{7}\left(\frac{y}{y_0}\right)^4 - 2\left(\frac{y}{y_0}\right)^6 + \left(\frac{y}{y_0}\right)^8\right]. \qquad \ldots\ldots(16)$$

The shape of the $\mathrm{OPD}_{\mathrm{B}}^*$ curve is shown in figure 5. In a similar way we find

$$R = 1 - \tfrac{1}{2}a_4{}^2\left(\frac{2\pi}{\lambda_0}\right)^2\left[\frac{1}{44100}\right], \qquad \ldots\ldots(17)$$

assuming that the OPD residues do not greatly exceed the Rayleigh limit. To find the maximum allowable tolerance, if we still use the 11 per cent amplitude reduction limit, for $R \geqslant 1 - 0{\cdot}11$, $|a_4| \leqslant 15\lambda_0$. The corresponding Rayleigh limit is shown by the two horizontal dotted lines for a_4 at full tolerance. Again we note there are two very narrow zones near the axis and the extreme margin whose $\mathrm{OPD}_{\mathrm{B}}^*$ values exceed $\frac{1}{4}\lambda_0$, but they would not cause serious error on using the simplified expression (9). Hence we come to the second conclusion:

(2) In a well-designed optical system, where the axial spherical aberrations are controllable up to the second order, the tertiary aberration, when put in terms of $\mathrm{OPD}_{\mathrm{B}}$, should not exceed $15\lambda_0$.

In a well-known treatment by Prof. A. E. Conrady (1923), the tolerance for tertiary aberration is given a limiting value of $8\frac{1}{2}\lambda_0$ as the maximum which

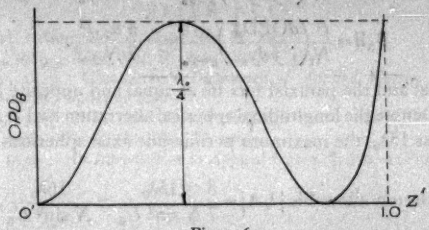


Figure 6.

can be allowed (under stated conditions) without transgression of the Rayleigh quarter-wave-length limit. In this case the OPD curve appears as shown in figure 6. His treatment is based on the idea of minimum optical path-difference. The present treatment shows that the tolerance may be practically doubled. Although we must acknowledge that the assumption of the presence of tertiary aberration with negligible higher-order terms is only a rough approximation, which, from the practical point of view, we must not regard as absolutely trustworthy, a realization of the true maximum permissible tolerance would nevertheless be a good guide.

In the scheme of design suggested by Professor Conrady the control of the primary and secondary aberration is obtained by bringing the paraxial $\frac{10}{11}$ zonal, and $\frac{10}{11}\frac{1}{\sqrt{2}}$ zonal foci to the same point. In the present scheme, in order to allow a greater tolerance, the earlier method has to be modified to some extent.

In equation (16), to find the geometrical ray intersections of different zones, we differentiate it:

$$\frac{d\mathrm{OPD}_B^*}{dy} = 4\frac{a_4}{y_0}\left[-\frac{1}{7}\left(\frac{y}{y_0}\right)+\frac{9}{7}\left(\frac{y}{y_0}\right)^3-3\left(\frac{y}{y_0}\right)^5+2\left(\frac{y}{y_0}\right)^7\right].$$

Firstly, $\frac{d\mathrm{OPD}_B}{dy}$ vanishes at $y=0{\cdot}416$, $y=0{\cdot}707$, $y=0{\cdot}909$ zones. Therefore, if we make ray tracings through these zones, these three rays will meet at the same focus, and this is the best focus. The corresponding OPD_B values are $\mathrm{OPD}_{B_{0\cdot416}}=\mathrm{OPD}_{B_{0\cdot909}}=-\frac{a_4}{45}$, $\mathrm{OPD}_{B_{0\cdot707}}=-\frac{a_4}{112}$. When $|a_4|=15\lambda_0$, $\mathrm{OPD}_{B_{0\cdot416}}=\mathrm{OPD}_{B_{0\cdot909}}=\lambda_0/3$. Secondly, at the axis and at the extreme margin,

$$\left(\frac{d\mathrm{OPD}_B^*}{y\,dy}\right)_{y=0}=-\frac{4}{7}\frac{a_4}{y_0^2},\qquad \left(\frac{d\mathrm{OPD}_B^*}{y\,dy}\right)_{y=y_0}=\frac{4}{7}\frac{a_4}{y_0^2}.$$

Denote the marginal focus by B_M and the paraxial by B_p. The distances B_MB and $B\,B$ are given by the expressions

$$B_MB=\frac{l^2}{N}\left(\frac{d\mathrm{OPD}_B^*}{y\,dy}\right)_{y=y_0}=+\frac{4}{7}\frac{a_4}{N}\frac{l^2}{y_0^2},$$

$$B_pB=\frac{l^2}{N}\left(\frac{d\mathrm{OPD}^*}{y\,dy}\right)_{y=y_0}=-\frac{4}{7}\frac{a_4}{N}\frac{l^2}{y_0^2}.$$

Thus the marginal and the paraxial foci lie at equal and opposite distances from the best focus. Denote the longitudinal spherical aberration by LA. $LA=B\,B_M$. By allowing $|a_4|$ as $15\lambda_0$, the maximum permissible axial spherical-aberration LA becomes

$$\text{Maximum allowable}\,|LA|=\frac{8}{7}\frac{15\lambda_0}{N\sin^2U_0}\simeq\frac{16\lambda_0}{N\sin^2U_0},\quad\ldots\ldots(18)$$

which is 16 times the axial focal range of the system. Hence we have two more conditions:

(3) The best focus lies midway between the paraxial and the marginal foci. It is the point where the 0·416, 0·707, and 0·909 zonal rays meet.

(4) The maximum allowable longitudinal spherical-aberration is 16 times the focal range of the system.

In practical designing work, for the final achievement to be in accordance with expression (15), we may either:

(*a*) bring the 0·416, 0·707, and 0·909 zonal rays to the same focus, the tolerance being guarded against by comparing the value of LA obtained with expression (18), or

(*b*) bring the 0·416 and 0·909 zonal rays to the same focus, and in addition to that, $\mathrm{OPD}_{B_{0\cdot416}}=\mathrm{OPD}_{B_{0\cdot909}}$, the performance of the system then being checked by observing $|\mathrm{OPD}_{B_{0\cdot416}}|=|\mathrm{OPD}_{B_{0\cdot902}}|\leqslant\lambda/3$.

It should be remembered that in a system where tertiary aberration plays a sensible part, aberrations of still higher orders are likely to be present and may even become very big at the extreme margin. Referring to figure 5, the OPD curve usually tends to become steeper at the extreme marginal zone than we have here assumed. This will cause the marginal focus to shift farther away from the best focus, thus giving a greater LA. Therefore the restriction of LA below a certain limit is naturally a safeguard against higher-order aberrations. In the case of method (*b*) being used, as just suggested, it is advisable still to trace the marginal and paraxial rays to see what the value of LA looks like.

It may also be noted that the so-called focal shift as a controllable factor occurs only in the theory, but not in the practical design. Once the optical system is properly designed, the judgement of the best focus is left to the observer.

§ 6. ACKNOWLEDGEMENTS

In conclusion, the author wishes to express his thanks to Professor L. C. Martin, who carefully read the manuscript and pointed out several corrections which should be made. He has also the pleasure of thanking the authorities of the Imperial College of Science and Technology for allowing him to publish this note.

REFERENCES

BUXTON, A., 1921. *Mon. Not. R. Astr. Soc.* **81**, 547.
CONRADY, A. E., 1919. *Mon. Not. R. Astr. Soc.* **79**, 575.
CONRADY, A. E., 1923. Glazebrook's *Dictionary of Applied Physics*, **4**, 215 (under the title *Microscope, Optics of the*).
CONRADY, A. E., 1926. *Proc. Opt. Conv.* Part 2, 830.
MARTIN, L. C., 1926. *Trans. Opt. Soc.* **27**, 249.
MARTIN, L. C., 1930. *An Introduction to Applied Optics*, **1**, 121–6 (Pitman and Co.).

LONGITUDINAL ELECTROMAGNETIC WAVES BETWEEN PARALLEL PLATES

By V. C. A. FERRARO and H. T. FLINT,
King's College, London*
* Now at H. H. Wills Physical Laboratory, Bristol

Received 31 *October* 1940

ABSTRACT. The problem considered is an example of the propagation of electromagnetic waves between perfectly conducting parallel plates, where it is possible to see how the conductivity of the dielectric affects the propagation. In addition, the simplicity of the functions involved in the solution of the problem enables its physical aspect to be readily appreciated.

§1. INTRODUCTION

The first problem of the type now to be considered was studied by Rayleigh (1897), who examined the case of the propagation of electromagnetic waves in the hollow space bounded by a conducting cylinder. He discovered the interesting results that the waves transmitted had components of intensity in the direction of propagation and that a " cut-off " frequency existed, that is to say, that there is a limiting frequency below which transmission along the tube is impossible. The magnitude of this frequency corresponds to a wave-length in the unbounded medium of the order of magnitude of the diameter of the cylinder.

At the time when this problem was first considered there was no question of experimental verification of Rayleigh's deductions, but the recent developments in the methods of generation of very short electromagnetic waves—that is, of wave-lengths in air of 10 cm. and less—have made these results of practical importance. The study of the propagation of these waves in bounded media has thus received much recent attention, particularly by a number of American writers, who have considered in detail, both experimentally and theoretically, cases of hollow tubes of rectangular and circular cross-sections (Southworth, 1936; Carson, Mead and Schelkunoff, 1936; Barrow, 1936; Schelkunoff, 1937).

The case of propagation between two infinitely extended parallel conducting plates will be considered here. The conditions will be idealized by making the plates perfect conductors, but the effect of conductivity in the dielectric will be examined. It appears worth while to study this case for two reasons: it offers a simple example of wave propagation in hollow guides, the solution being obtained in terms of very simple mathematical expressions, so that the physical

aspect of the problem is not obscured, and it enables one to see how the conductivity of the medium affects the propagation.

The problem is approached by the general method of introducing the vector and scalar potential, and follows the method used by Hondros and Debye (1910) in their study of transmission on a dielectric wire.

§ 2. THE GENERAL APPROACH TO THE PROBLEM

The straightforward method of approach to a solution is by means of the electromagnetic scalar and vector potentials, ϕ and A respectively. Let the medium have the coefficient of conduction σ, expressed in e.m.u. With the usual vector notation, Maxwell's equations are:

$$\frac{K}{c}\dot{\boldsymbol{E}}+4\pi\sigma\boldsymbol{E}=\operatorname{curl}\boldsymbol{H}, \qquad \ldots\ldots(1)$$

$$\frac{\mu}{c}\dot{\boldsymbol{H}}=-\operatorname{curl}\boldsymbol{E}, \qquad \ldots\ldots(2)$$

$$\operatorname{div}\boldsymbol{E}=0, \qquad \ldots\ldots(3)$$

$$\operatorname{div}\boldsymbol{H}=0. \qquad \ldots\ldots(4)$$

Equation (3) implies that the medium is uncharged, the value of the electric density being zero. K denotes the dielectric constant and μ the permeability of the medium.

Introducing the scalar and vector potentials, the electric and magnetic field vectors are given by

$$\mu\boldsymbol{H}=\operatorname{curl}\boldsymbol{A}, \quad \boldsymbol{E}=-\frac{1}{c}\dot{\boldsymbol{A}}-\operatorname{grad}\phi. \qquad \ldots\ldots(5)$$

It follows from (1) and (5) that

$$\left(\frac{K}{c}\frac{\partial}{\partial t}+4\pi\sigma c\right)\boldsymbol{E}=\operatorname{curl}\boldsymbol{H}=\operatorname{curl}^2\frac{\boldsymbol{A}}{\mu}=\operatorname{grad}\left(\operatorname{div}\frac{\boldsymbol{A}}{\mu}\right)-\nabla^2\left(\frac{\boldsymbol{A}}{\mu}\right). \qquad \ldots\ldots(6)$$

Assuming that all the quantities are periodic in t, which occurs in the factor $e^{i\omega t}$, and writing

$$K'=K-4\pi i\frac{\sigma c^2}{\omega}, \qquad \ldots\ldots(7)$$

it follows that

$$\frac{K'i\omega}{c}\boldsymbol{E}=\operatorname{grad}\left(\operatorname{div}\frac{\boldsymbol{A}}{\mu}\right)-\nabla^2\left(\frac{\boldsymbol{A}}{\mu}\right). \qquad \ldots\ldots(8)$$

Also from (5)

$$\frac{K'i\omega}{c}\boldsymbol{E}=-\frac{K'i\omega}{c^2}\boldsymbol{A}-\operatorname{grad}\frac{K'i\omega}{c}\phi. \qquad \ldots\ldots(9)$$

Without loss of generality ϕ may be chosen so that

$$\operatorname{div}(\boldsymbol{A}/\mu)+(K'i\omega/c)\phi=0, \qquad \ldots\ldots(10)$$

whence

$$\nabla^2\left(\frac{\boldsymbol{A}}{\mu}\right)=\frac{K'i\omega}{c^2}\boldsymbol{A}. \qquad \ldots\ldots(11)$$

Writing

$$\frac{\boldsymbol{A}}{\mu}=\frac{K'i\omega}{c}\boldsymbol{P}, \quad \text{so that } \phi=-\operatorname{div}\boldsymbol{P}, \qquad \ldots\ldots(12)$$

it follows that

$$\nabla^2\boldsymbol{P}=\frac{\dot{\boldsymbol{A}}}{c}=-\frac{\mu K'\omega^2}{c^2}\boldsymbol{P} \qquad \ldots\ldots(13)$$

The solution of the problem thus depends on the determination of the vector $\boldsymbol{P}$, the field intensities being determined by

$$\boldsymbol{E} = \text{grad (div } \boldsymbol{P}) - \frac{\mu K' i\omega^2}{c^2}\boldsymbol{P}, \qquad \ldots\ldots(14)$$

$$\boldsymbol{H} = \frac{K' i\omega}{c}\,\text{curl } \boldsymbol{P}. \qquad \ldots\ldots(15)$$

The direction of propagation of the wave will be supposed to be along the positive axis of z, and this variable will occur only in the factor e^{-imz}. If m is real, the wave travels without damping along the direction of z, but if m is imaginary there is no wave and the oscillations die away as z increases; e^{-imz} then takes the form e^{-az}, where a is a real positive number. The case of a negative value for a is excluded since this would mean unlimited increase of the quantities concerned as z increases. Finally, the case of a complex value of m means that a wave occurs with exponential decay of amplitude as the z-axis is traversed.

The quantity $\boldsymbol{P}$ thus contains the factor $e^{i(\omega t - mz)}$, and it will be replaced by

$$\boldsymbol{P} = \boldsymbol{T} e^{i(\omega t - mz)}, \qquad \ldots\ldots(16)$$

where $\boldsymbol{T}$ is a vector depending upon the variables x and y. Equation (13) can now be replaced by

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\boldsymbol{T} + \left(\frac{\mu\omega^2 K'}{c^2} - m^2\right)\boldsymbol{T} = 0 \qquad \ldots\ldots(17)$$

and with

$$\chi^2 = \frac{\mu\omega^2 K'}{c^2} - m^2 \qquad \ldots\ldots(18)$$

this becomes

$$\frac{\partial^2 \boldsymbol{T}}{\partial x^2} + \frac{\partial^2 \boldsymbol{T}}{\partial y^2} = -\chi^2 \boldsymbol{T}. \qquad \ldots\ldots(19)$$

In order to solve any particular problem the solution of this equation which satisfies the appropriate boundary conditions is required.

The values of the field intensities may be found from equations (14) and (15). Omitting for convenience the factor $e^{i(\omega t - mz)}$, by which all the functions have to be multiplied, it is found that

$$E_x = \frac{\partial}{\partial x}\left(\frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y} - imT_z\right) + \frac{\mu K'\omega^2}{c^2}T_x, \qquad \ldots\ldots(20)$$

$$E_y = \frac{\partial}{\partial y}\left(\frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y} - imT_z\right) + \frac{\mu K'\omega^2}{c^2}T_y, \qquad \ldots\ldots(21)$$

$$E_z = -im\left(\frac{\partial T_x}{\partial x} + \frac{\partial T_y}{\partial y}\right) + \chi^2 T_z, \qquad \ldots\ldots(22)$$

$$H_x = \frac{K' i\omega}{c}\left(\frac{\partial T_z}{\partial y} + imT_y\right), \qquad \ldots\ldots(23)$$

$$H_y = -\frac{K' i\omega}{c}\left(\frac{\partial T_z}{\partial x} + imT_x\right), \qquad \ldots\ldots(24)$$

$$H_z = \frac{K' i\omega}{c}\left(\frac{\partial T_y}{\partial x} - \frac{\partial T_x}{\partial y}\right). \qquad \ldots\ldots(25)$$

The x- and y-components may be conveniently expressed in terms of E_z and H_z as follows, provided that these components exist:

$$E_x = -\frac{i}{\chi^2}\left(m\frac{\partial E_z}{\partial x} + \frac{\mu\omega}{c}\frac{\partial H_z}{\partial y}\right), \qquad \ldots\ldots(26)$$

$$E_y = -\frac{i}{\chi^2}\left(m\frac{\partial E_z}{\partial y} - \frac{\mu\omega}{c}\frac{\partial H_z}{\partial x}\right), \qquad \ldots\ldots(27)$$

$$H_x = \frac{i}{\chi^2}\left(\frac{K'\omega}{c}\frac{\partial E_z}{\partial y} - m\frac{\partial H_z}{\partial x}\right), \qquad \ldots\ldots(28)$$

$$H_y = -\frac{i}{\chi^2}\left(\frac{K'\omega}{c}\frac{\partial E_z}{\partial x} + m\frac{\partial H_z}{\partial y}\right). \qquad \ldots\ldots(29)$$

A wave with a longitudinal component of electric intensity and with no corresponding magnetic component is described as an E-wave or as a transverse *magnetic* wave. Likewise if E_y vanishes and H_z exists, the wave is an H-wave or a transverse *electric* wave.

The formulae (26) to (29) show that the general case may be regarded as the superposition of the two types of waves. It is possible by experiment to produce both E- and H-waves, and the properties of these two types can be considered separately.

§ 3. THE TRANSVERSE MAGNETIC OR E-WAVE

In this case $H_z = 0$, so that by equation (25)

$$\frac{\partial T_y}{\partial x} = \frac{\partial T_x}{\partial y}.$$

A problem in which this equation is satisfied most simply will be considered, viz., that in which T_x and T_y vanish. The vector $\boldsymbol{T}$ now has only the component T_z, and by equation (22)

$$E_z = \chi^2 T_z. \qquad \ldots\ldots(30)$$

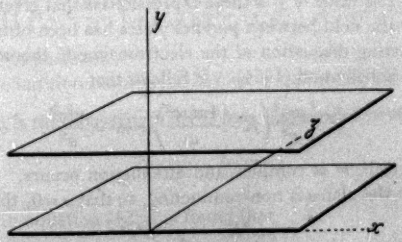


Figure 1. Conducting plates.

Consider two conducting plates, of which the lower occupies the zx-plane (figure 1), the upper plate occupying the plane $y = a$, where a is their separation.

Suppose that the plates are perfect conductors. This means that the electric intensity in each plate vanishes and, since the tangential components of the intensity are continuous at the surface of separation of two media, it follows that E_x and E_z both vanish in the dielectric at $y=0$ and at $y=a$.

The form of T_z is simplified in this case, for it is a function of y only.

Equation (19) becomes

$$d^2T_z/dy^2 = -\chi^2 T_z, \qquad \ldots\ldots(31)$$

and the solution is

$$\left.\begin{aligned} T_z &= Ae^{i\chi y} + Be^{-i\chi y} \\ \text{and} \quad E_z &= \chi^2(Ae^{i\chi y} + Be^{-i\chi y}). \end{aligned}\right\} \qquad \ldots\ldots(32)$$

Since, in this case, it is postulated that H_z is zero,

$$E_x = 0, \quad E_y = m\chi(Ae^{i\chi y} - Be^{-i\chi y}), \qquad \ldots\ldots(33)$$

$$H_z = -\frac{K'\omega\chi}{c}(Ae^{i\chi y} - Be^{-i\chi y}), \quad H_y = 0. \qquad \ldots\ldots(34)$$

One boundary condition requires $E_z=0$ at $y=0$, i. e. $A+B=0$. Thus

$$E_z = \chi^2 A(e^{i\chi y} - e^{-i\chi y}), \qquad \ldots\ldots(35)$$

and since a second boundary condition requires $E_y=0$ at $y=a$, it follows that

$$e^{i\chi a} - e^{-i\chi a} = 2i \sin \chi a = 0.$$

Thus

$$\chi a = n\pi, \qquad \ldots\ldots(36)$$

where n is zero or an integer.

The zero value applies to the trivial case when there is no field. From equations (26) to (29) it follows that the expressions for the field components are

$$E_x = 0, \quad E_y = m\chi A(e^{i\chi y} + e^{-i\chi y}), \quad E_z = \chi^2 A(e^{i\chi y} - e^{-i\chi y}), \qquad \ldots\ldots(37)$$

$$H_x = -\frac{K'\omega\chi}{c} A(e^{i\chi y} + e^{-i\chi y}), \quad H_y = 0, \quad H_z = 0. \qquad \ldots\ldots(38)$$

The complete expressions for the components are obtained by multiplying these expressions by the factor $e^{i(\omega t - mz)}$ and the real or imaginary parts taken as required in actual cases. The value of χ in these expressions is that given by equation (36). The solution for the field between parallel plates has been obtained as a limiting case in an interesting discussion of the electromagnetic theory of coaxial transmission lines by Schelkunoff (1934). It follows that

$$\chi^2 = \frac{\mu\omega^2}{c^2}\left(K - \frac{4\pi i\sigma c^2}{\omega}\right) - m^2 = \frac{n^2\pi^2}{a^2}. \qquad \ldots\ldots(39)$$

If σ does not vanish, m is complex and attenuation occurs. Suppose that the medium between the plates is non-conducting, so that $\sigma=0$, then

$$\frac{\mu\omega^2 K}{c^2} - m^2 = \frac{n^2\pi^2}{a^2}.$$

In order that a wave may exist m must be real, and consequently

$$\mu\omega^2 K/c^2 > n^2\pi^2/a^2.$$

The velocity of the wave in the free medium is $V=c/\sqrt{(\mu K)}$, so that this inequality can be written

$$\omega/V>n\pi/a.$$

For the value $n=1$

$$\omega>\pi V/a,$$

$$2\pi\nu>\pi V/a.$$

The lowest value of ν is

$$\nu_c=V/2a. \qquad \ldots\ldots(40)$$

This corresponds to a limiting wave-length in the free medium of magnitude $2a$.

For the nth node

$$\nu_c=nV/2a. \qquad \ldots\ldots(41)$$

There exists a " cut-off " frequency ν_c which has its lowest value $V/2a$, corresponding to the lowest mode with

$$E_z \propto \sin \pi y/a.$$

A similar discussion applies to the transverse electric or H-wave.

§ 4. THE CASE OF A CONDUCTING MEDIUM

It will be assumed that the boundaries are again perfect conductors, but that σ does not vanish for the medium.

In this case m is complex and will be written in the form $m=\beta-i\alpha$. In this case the factor $e^{i(\omega t-mz)}$ becomes $e^{-\alpha z}e^{i(\omega t-\beta z)}$, so that α denotes the attenuation constant and $\beta=2\pi/\lambda$, where λ is the wave-length within the bounded medium.

From equation (39),

$$\alpha^2-\beta^2=\chi^2-\mu\omega^2K/c^2 \quad \text{and} \quad 2\alpha\beta=4\pi\mu\omega\sigma. \qquad \ldots\ldots(42)$$

Thus

$$\alpha=\mu\omega\sigma\lambda \qquad \ldots\ldots(43)$$

and

$$\chi^2=\frac{\mu K\omega^2}{c^2}-\beta^2+\frac{4\pi^2\mu^2\omega^2\sigma^2}{\beta^2}. \qquad \ldots\ldots(44)$$

For the present case a real value of β is sought. Such a value is always possible, for a solution of equation (44) is

$$\beta^2=-\tfrac{1}{2}\left(\chi^2-\frac{\mu K\omega^2}{c^2}\right)+\tfrac{1}{2}\sqrt{\left(\chi^2-\frac{\mu K\omega^2}{c^2}\right)^2+16\pi^2\omega^2\sigma^2\mu^2},$$

and this value of β^2 is positive.

The boundary condition requires, as before, $\chi=n\pi/a$. If the waves are long, i.e. $\lambda_0\gg 2a$, where λ_0 is the wave-length in the free, unbounded medium, $\chi^2\gg\frac{\mu K}{c^2}\omega^2$, and hence

$$\beta\approx\frac{2\pi\mu\sigma\omega}{\chi}=\frac{2\sigma a\mu\omega}{n}.$$

After substitution in equation (42) it is found that

$$\alpha=n\pi/a, \qquad \ldots\ldots(45)$$

so that the attenuation factor is $e^{-n\pi z/a}$, the amplitude being reduced to $1/e$ of its value in a distance $a/n\pi$.

It has been assumed that $\lambda_0\gg 2a$, so that the wave-length is much greater than

$2a/n$, and this reduction therefore takes place in a small fraction of the wavelength. This means that the wave can be perceived only in a distance small compared with the separation of the plates, and thus there is in effect a cut-off of long waves.

§ 5. THE RELATION BETWEEN WAVE NUMBER AND FREQUENCY

Write

$$N=1/\lambda,\ \nu=\omega/2\pi,\ V=c/\sqrt{\mu K},$$

so that equation (44) becomes

$$\frac{1}{V^2}+\frac{\sigma^2\mu^2}{N^2}-\frac{N^2}{\nu^2}-\frac{\chi^2}{4\pi^2\nu^2}=0.$$

Write $A=\sigma^2\mu^2,\quad B=\chi^2/4\pi,\quad x=N^2,\quad y=\nu^2,$

then $$xy+AV^2y-V^2x^2-BV^2x=0. \qquad \ldots\ldots(46)$$

This is the equation of a hyperbola whose asymptotes are

$$x+AV^2=0 \quad \text{and} \quad y-V^2x-V^2(B-AV^2)=0. \qquad \ldots\ldots(47)$$

Three cases can be considered, according as $B>AV^2$, $B<AV^2$ or $B=AV^2$.

Case I. $B>AV^2$ (i.e. $\chi>2\pi\sigma\mu$)

This is the case in practice for very badly conducting media, since with $\sigma=10^{-12}$ e.m.u. the condition for the lowest mode is $a<\frac{1}{2}\times 10^{12}$ cm.

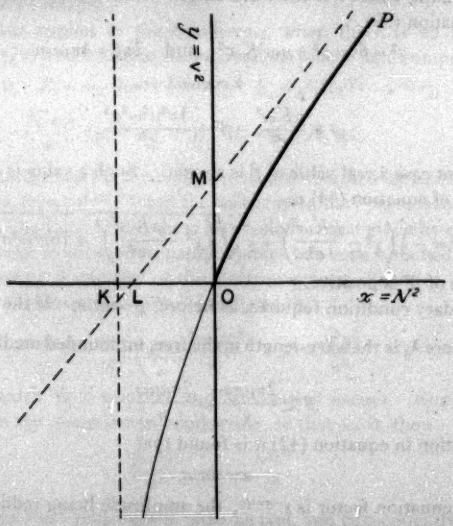


Figure 2. Case I. $B>AV^2$.

Equation (46) is represented in figure 2, the part OP of the curve being of physical significance and the other branch of the hyperbola being omitted.

$$\text{OK} = AV^2 = \sigma^2\mu^2V^2,$$

and corresponds to a wave number $N_0 = \sigma\mu V$.

$$\text{OL} = B - AV^2 = \frac{\chi^2}{4\pi^2} - \sigma^2\mu^2V^2.$$

L lies to the left of O since $B > AV^2$.

From (43) the value of the attenuation constant is

$$\alpha = 2\pi\mu\sigma\nu/N = 2\pi\mu\sigma\sqrt{y/x}. \qquad \ldots\ldots(48)$$

Thus

$$\frac{y}{x} = V^2 \cdot \frac{1 + B/x}{1 + AV^2/x}.$$

As the wave-length diminishes, i.e. $x \to \infty$, this ratio approaches the constant value V^2 and the constant $\alpha \to 2\pi\mu\sigma V$. On the other hand, as the wave-length increases, i.e. $x \to 0$, the ratio approaches the value $B/A = \chi^2/4\pi^2\sigma^2\mu^2$ and $\alpha \to \chi$. This is the case already considered for long waves.

The condition $B > AV^2$ means $\chi > 2\pi\mu\sigma V$, so that the attenuation for long waves is greater than for short ones.

Case II. $B < AV^2$ (i.e. $\chi < 2\pi\sigma\mu$)

Similar conclusions to those of the preceding case can be made, but the order of magnitude of the attenuation constants is reversed. For long waves the value is less than for short waves. The graph in this case is shown in figure 3.

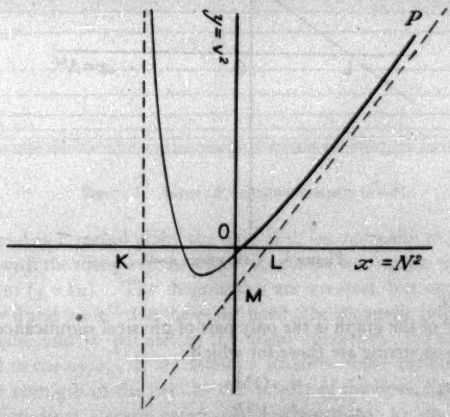


Figure 3. Case II. $B < AV^2$.

Case III. $B = AV^2$ (i.e. $\chi = 2\pi\sigma\mu$)

Equation (46) becomes in this case

$$(x + AV^2)(y - V^2x) = 0.$$

Thus the part of the graph of physical significance is the straight line,

$$y = V^2 x. \qquad \ldots\ldots(49)$$

In this case the attenuation is the same for all wave-lengths:

$$\alpha = 2\pi\sigma\mu V = n\pi/a.$$

Case of $\sigma = 0$.

It is interesting to consider this case in connection with the foregoing general argument.

In this case the constant A is zero and equation (46) becomes

$$x(y - V^2 x - BV^2) = 0,$$

and the graph for the relation between frequency and wave number becomes the straight line

$$y = V^2(x + B).$$

In figure 4, $\mathrm{OL} = B$, $\mathrm{OM} = V^2 B$.

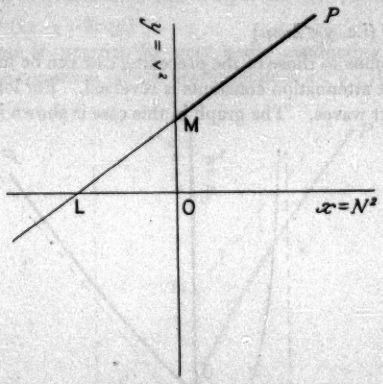


Figure 4. Case where $\sigma = 0$.

Since $\sigma = 0$, $\alpha = 0$.

The part MP of the graph is the only part of physical significance. Thus the only frequencies occurring are those for which

$$\begin{aligned} & y > \mathrm{OM}, \\ \text{i.e.} \quad & \nu^2 > V^2 B, \\ & \nu^2 > V^2 \chi^2/4\pi^2, \\ & \nu > V\chi/2\pi, \\ \text{or} \quad & \lambda_0 < 2a/n, \end{aligned}$$

which is the result obtained before.

Numerical example (Case I)

Let the medium between two parallel plates be quartz, 500 cm. thick. For quartz the specific resistance is of the order 10^8 ohms per cm. cube, i.e. 10^{17} e.m.u., so that $\sigma = 10^{-17}$. For long waves $\beta = 2\pi/\lambda = 2a\sigma\mu\omega$ and the phase velocity of the wave between the plates is $V' = \omega\lambda/2\pi$. Thus if $\mu = 1$, $V' = 10^{14}$ cm. per sec. This being a phase velocity, the fact that it is greater than the velocity of light is not surprising.

In a distance of 1000 cm., i.e. twice the separation of the plates, the amplitude is reduced by a factor of e^{-6}, i.e. practically complete absorption has occurred.

§ 6. DIAGRAMS OF THE FIELD DISTRIBUTION

A conception of the physical character of the field of the *E*-wave may be obtained from a diagrammatic representation of the lines of force.

Consider the case in which $n = 1$. The value of the component of electric intensity in the plane perpendicular to the direction of propagation is proportional

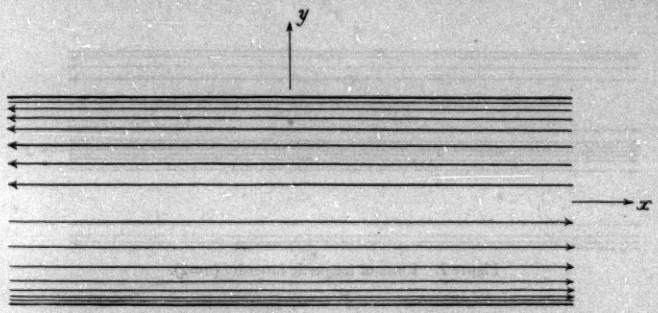


Figure 5. Lines of magnetic intensity ($n=1$).

to $\cos \pi y/a$ and is directed along the axis of y, i.e. normally to the plates. The value varies with the distance from one of the plates, having the value zero midway between them ($y = \frac{1}{2}a$). The magnitudes are greatest, but opposite in sign, at the plates $y = 0$ and $y = a$. On the other hand, the magnetic intensity lies wholly in the xy-plane and is parallel to the plate surfaces. Its magnitude is also proportional to $\cos \pi y/a$. If we attempt to show these variations in a diagram denoting the strength of the field by the density of the lines, figure 5 is obtained.

The longitudinal electric intensity is proportional to $\sin \pi y/a$ and has its greatest value midway between the plates, vanishing at the plates themselves.

The variation in intensity can be represented by drawing lines in the direction of propagation with a density proportional to the intensity at various points between the plates, as in figure 6.

It is interesting to repeat this process for the second possible mode of the propagation, viz. that for $n=2$. In this case the amplitudes are proportional to $\cos 2\pi y/a$ and $\sin 2\pi y/a$.

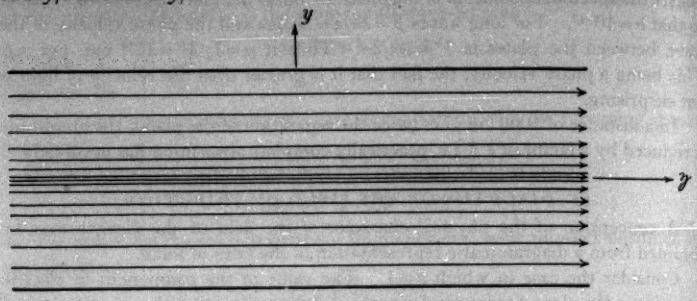


Figure 6. Lines of longitudinal electric intensity ($n=1$).

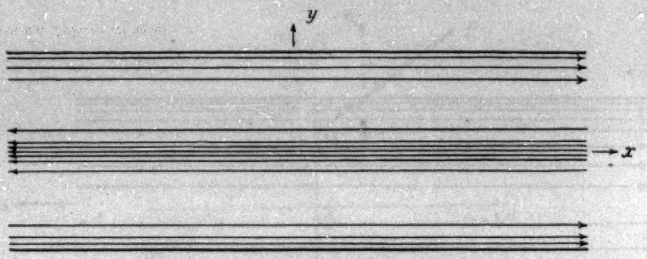


Figure 7. Lines of magnetic intensity ($n=2$).

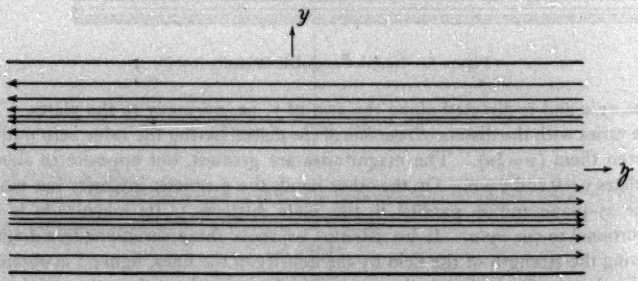


Figure 8. Lines of longitudinal electric intensity ($n=2$).

The electric intensity normal to the plates varies in proportion to $\cos 2\pi y/a$ and has the greatest value at one plate, falling through the value 0 at $y=a/4$. From this point its sign changes and the greatest magnitude is again attained at

$y=\frac{1}{2}a$, midway between the plates but with sign reversed. A similar sequence of changes occurs in the upper half of the field, the maximum value being again reached at the upper plate.

In the same way the diagram for the lines of magnetic intensity in this mode is obtained, and illustrated in figure 7.

In the case of the longitudinal component of electric intensity the variation is proportional to $\sin 2\pi y/a$, and is represented in figure 8 on the same principle as that used to represent the corresponding case of the first mode.

REFERENCES

BARROW, 1936. *Proc. Inst. Radio Engrs*, **24**, 1298.
CARSON, MEAD and SCHELKUNOFF, 1936. *Bell Syst. Tech. J.* **15**, 310.
HONDROS and DEBYE, 1910. *Ann. Phys., Lpz.*, **36**, 465.
RAYLEIGH, 1897. *Phil. Mag.* **43**, 125.
SCHELKUNOFF, 1934. *Bell Syst. Tech. J.* **13**, 541.
SCHELKUNOFF, 1937. *Proc. Inst. Radio Engrs*, **25**, 1457.
SOUTHWORTH, 1936. *Bell Syst. Tech. J.* **15**, 284.

ON SURFACE ENERGY AND SURFACE ENTROPY

By ALLAN FERGUSON and ERIC J. IRONS,
Queen Mary College (now at the Cavendish Laboratory, Cambridge)

Received 16 *November* 1940

ABSTRACT. It is shown that total molar surface energy (ΔE) and molar surface entropy (ΔS), when defined by the equations

$$\Delta E=\mu-T\partial\gamma/\partial T \quad \text{and} \quad \Delta S=-\partial\mu/\partial T,$$

where μ is free molar surface energy, are constants independent of temperature, and are connected by the relation

$$T_c \,.\, \Delta S=\Delta E,$$

where T_c is the critical temperature.

Further, ΔE, defined in this way, is identified with E_0, the zero value of the molar surface energy of the supercooled liquid, when molar surface energy is defined in the usual manner.

In what follows, we shall consider only the work term $\gamma\, da$, where γ stands for surface tension (or free surface energy per unit area), and da for an increment of area. Then, where λ stands for total surface energy per unit area, we have, as is well known,

$$\lambda=\gamma-T\left(\frac{\partial\gamma}{\partial T}\right)_a,$$

where the subscript a may be dropped, inasmuch as, save in very exceptional circumstances, γ is independent of area.

Also, where S stands for entropy,

$$S=-\left(\frac{\partial\gamma}{\partial T}\right)_a.$$

It is useful, however, to consider the amounts of these quantities associated with the molar surface $(Mv)^{2/3}$ or $(M/\rho)^{2/3}$, where M stands for molecular weight and v for specific volume. As is commonly understood, if μ is the molar free surface energy, defined by

$$\mu=\gamma(M/\rho)^{2/3},$$

then the total molar surface energy (ν) is defined as

$$\nu=\lambda(M/\rho)^{2/3}.$$

Campbell and Eley (1940), who have recently published an interesting note on the topic of surface energy and entropy, adopt a slightly different definition of

total molar surface energy. We shall use their notation in the argument which follows. ΔS, the molar surface entropy, being given by $\Delta S = -(\partial\mu/\partial T)$, the molar surface energy, ΔE, is given by $\Delta E = \mu - T(\partial\mu/\partial T)$. They calculate values of ΔS and of ΔE from the known values of surface tension and density (and their temperature coefficients) at 25° C. ($T = 298°$ K.), and find that for a large number of pure liquids $T.\Delta S$ is proportional to ΔE.

Now, we know that, *for unassociated liquids*, the equation

$$\gamma = \gamma_0(1 - bT)^n,$$

with n taken as 1·2, represents the variation of surface tension with temperature very closely over the whole range of existence of the liquid, from freezing point to critical temperature. The constant b is obviously the reciprocal of the critical temperature, T_c. If it is desired to calculate T_c accurately from observations

Table 1

Substance	$T.\Delta S$ ergs	ΔE ergs	T_c (calc.) ° K. ($=\Delta E/\Delta S$)	T_c (exper.) ° K. (from tables)
1. *m*-Xylene	677	1397	615	622
2. *n*-Propylbenzene	648	1417	652	639
4. *n*-Propyl chloride	655	1073	488	503
7. Methyl alcohol	240	503	625	513
11. Acetic anhydride	777	1233	473	569
13. Acetic acid	325	732	671	595
14. Propionic acid	398	863	646	613
16. Ethyl acetate	630	1123	531	523
17. Ethyl propionate	643	1203	558	546
20. Acetonitrile	435	835	572	548
21. Di-isobutyl	573	1153	600	550
22. Methylene chloride	623	1162	556	518
25. Allyl chloride	573	997	518	514
31. Allyl alcohol	308	730	706	545
33. *n*-Butyric acid	602	1135	562	628
34. *n*-Valeric acid	570	1182	618	652
35. Isovaleric acid	465	1043	668	634
36. Ethyl formate	692	1123	484	508
37. *n*-Ethyl butyrate	878	1498	509	566
38. Tri-ethylamine	698	1240	529	535
A. Water	298	792	792	647
B. Ethyl alcohol	280	603	642	516
C. Isobutyl alcohol	375	833	662	538

We are indebted to Messrs. Campbell and Eley for the loan of the large-scale graph from which the diagram given in their paper was prepared. We have retained their numbering scheme for the substances and have read off, to about half a small square, their values of ΔE and $T.\Delta S$ ($T = 298°$ K.). From these values we have calculated, by means of the equation T_c (calc.) $= (\Delta E/\Delta S)$, the critical temperatures of those substances for which we have found the corresponding experimental values, T_c (exper.), in tables of constants.

of surface tension made over a limited range of temperature, the (usually slight) variation of n from the mean level 1·2 must be taken into account.

Macleod (1923) has shown that, over a similar range,

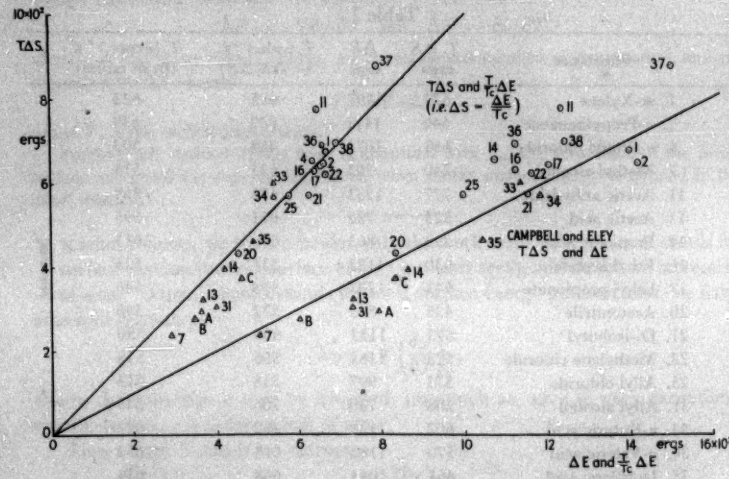
$$\gamma/(\rho_l-\rho_v)^p = C,$$

where Macleod takes p equal to 4. The symbol ρ_l stands for the orthobaric density of the liquid, ρ_v for that of the vapour. Here again, if high accuracy is desired, slight variations in the value of p must be taken into account.

At ordinary temperatures ρ_v is negligible, and we may write

$$\gamma/\rho^4 = C \quad \text{or} \quad \rho = \gamma^{1/4}/C^{1/4}.$$

{If, as Katayama (1916) has suggested, we define free and total molar surface energies as $\gamma[M/(\rho_l-\rho_v)]^{2/3}$ and $\lambda[M/(\rho_l-\rho_v)]^{2/3}$, we keep $(\rho_l-\rho_v)$ in Macleod's equation, and the investigation remains formally the same.}



We have, then,

$$\mu = \gamma(M/\rho)^{2/3} = \gamma M^{2/3} C^{1/6}/\gamma^{1/6} = C^{1/6} M^{2/3} \gamma^{5/6} = C^{1/6} M^{2/3} \gamma_0^{5/6}(1-bT).$$

Hence $$\Delta S = -\partial\mu/\partial T = bC^{1/6} M^{2/3} \gamma_0^{5/6}$$

and $$\Delta E = \mu - T\partial\mu/\partial T = C^{1/6} M^{2/3} \gamma_0^{5/6}.$$

We see, then, that

$$\Delta S = b \,.\, \Delta E, \quad \text{or} \quad \Delta E = T_c \,.\, \Delta S.$$

Table 1 shows values of T_c calculated from this equation, and compared with the experimental values. It must be emphasized that the equation $\Delta E = T_c \,.\, \Delta S$ applies only to unassociated liquids.

If ΔS is plotted against $\Delta E/T_c$, we should obtain a straight line equally inclined to the coordinate axes. This is shown in the accompanying figure, where, for

convenience of comparison, $T.\Delta S$ has been plotted against $T.\Delta E/T_c$. Campbell and Eley's graph of $T.\Delta S$ against ΔE is shown on the same scale; and values for alcohols and acids are indicated by triangles.

It will be noticed that the expressions deduced for ΔS and ΔE show that both ΔS and ΔE are independent of the temperature. If we define total molar surface energy as $\lambda[M/(\rho_l-\rho_v)]^{2/3}$, then it is easily seen that this quantity is a linear function of the temperature if we assume $n=1{\cdot}2$, $p=4$ (Katayama, 1916, and Ferguson and Kennedy, 1936). Defined by the expression $\lambda(M/\rho)^{2/3}$, molar surface energy is, as was shown by Ferguson and Miller (1934), a rather complicated function of the

Table 2

Substance	ΔE (Campbell and Eley, 1940)	E_0 (mean value derived from data of various experimenters by Ferguson and Kennedy, 1936)
m-Xylene	1397	1388
n-Propylbenzene	1417	1453
n-Propyl chloride	1073	1099
Ethyl acetate	1123	1170
Ethyl propionate	1203	1226
Ethyl formate	1123	1050

temperature. In fact, as has been previously shown (Ferguson and Kennedy, 1936), the value E_0 of the total molar surface energy of the supercooled liquid at the absolute zero is $C^{1/6}M^{2/3}\gamma^{5/6}$. Campbell and Eley's value of ΔE, when ΔE is defined as $\mu - T.\partial\mu/\partial T$, is the value of E_0 when E is defined in the more usual fashion.

Ferguson and Kennedy (1936) have given a fairly comprehensive table showing values of γ_0, n, p, and E_0 for some forty pure liquids. The values of E_0 calculated by them are compared in table 2 with the values of ΔE as directly calculated by Campbell and Eley.

REFERENCES

Campbell and Eley, 1940. *Trans. Faraday Soc.* **36**, 854.
Ferguson and Kennedy, 1936. *Trans. Faraday Soc.* **32**, 1474.
Ferguson and Miller, 1934. *Proc. Phys. Soc.* **46**, 140.
Katayama, 1916. *Sci. Rep. Tôhoku Imp. Univ.* **4**, 373.
Macleod, 1923. *Trans. Faraday Soc.* **9**, 38.

TRIBO-ELECTRICITY IN WOOL AND HAIR

By A. J. P. MARTIN, M.A., Ph.D.,
Wool Industries Research Association, Leeds

Received 14 *November* 1940

ABSTRACT. When wool fibres (or hairs) are rubbed together, the sign of the tribo-electric charges produced depends upon the sense of the rubbing. This effect is due to the cuticle cells of the fibre, which are piezo- and pyro-electric. Pyro-electricity is absent or weak in the cortical cells.

§1. INTRODUCTION

Wool and human hair and, presumably, the hair of any mammal—since all mammal hairs possess a similar histological structure—possess a directional tribo-electric effect. That is to say, when a fibre is rubbed against a similar fibre, it becomes charged, the sign of the charge acquired depending upon the direction of the rubbing.

Thus, if a number of similar fibres lie parallel, and all point in the same direction, i.e., all the root ends are together, then a fibre pulled out of the bundle by the root end carries a positive charge, a fibre pulled out by the tip a negative charge. As is to be expected, when a fibre lies in the opposite direction to its fellows no charge is produced whether it is pulled by the root or the tip. When the fibres are not parallel, the fibre drawn out carries a charge of the expected sign and of magnitude dependent on the angle between the fibres, a maximum at 0° and zero at 180°. The directional effect is still apparent when the fibre is rubbed against other materials, and though the sign of the charge may be the same, the magnitude varies with the direction of rubbing, being more positive when rubbed in the direction root to tip.

A fibre which has had the cuticle removed no longer possesses the directional effect, and becomes negatively charged on rubbing against undamaged fibres, showing ordinary tribo-electricity due to a difference between the cortical and the cuticle cells.

The explanation of these phenomena is that the scales or cuticle cells (but not the main part of the fibre, the cortical cells) are piezo-electric, so that when a shearing force in the direction of the tip of the fibre is applied to some small part of the surface of the fibre, that part becomes negatively charged.

The complementary positive charge may be localized either within the fibre or displaced towards the root and tip. Part at least of the negative charge leaks to whatever is applying the shearing force. When, therefore, a fibre is drawn

out of, say, a number of other fibres by its root, a negative charge will be transferred to its surroundings at all points at which it makes contact, and hence will itself carry a positive charge.

Now it is found that a piezo-electric charge is developed at each end of a bundle of fibres on compression, the root end becoming negative. On applying a shear along the length of a bundle of oriented wool fibres, a much smaller charge is developed on either side of the bundle. This indicates that the direction of the dipoles produced, assuming that this direction is always the same for any uniform stress, is at some slight angle to the length of the fibre. Since the scales can readily be seen under the microscope to lie at a small angle to the length of the fibre, it seems reasonable to assume at present that the line of separation of the charges coincides with the length of the scale.

The assumption that the direction of the separation of the charges is the same both in compression and other uniform stress is reasonable, as pyro-electricity can also be shown to be a property of the scales. Were there several directions in which piezo-electric charges could be developed, the scales would not be pyro-electric.

§2. EXPERIMENTAL

Preparation of oriented fibres. The fibres in a lock cut from the skin or picked from a fleece are, of course, already oriented, and the only preparation made was that of washing in soap and water and drying. They may also be conveniently prepared by drawing a length of top or a hank of untwisted yarn into a two- or three-inch length of rubber tube $^5/_{16}$″ internal diameter, and cutting close to the ends of the tube. On rolling this between a board and the bench for a few minutes, tufts will protrude at either end, and these tufts are composed entirely of root ends.

Detection of charges on the fibres. When single fibres were used they were held close to a wire on the terminal of a Lindemann electrometer, or of an electrometer triode in a pH-meter stabilized by a high resistance of pencil lines on paper between grid and earth, the latter being rather more convenient. When many fibres were used at once, a gold-leaf electroscope was amply sensitive enough.

To test rubbing at an angle to the length of the fibre, two microscope slides were covered at one end with oriented wool, the wool being attached on the back of the slide with sealing wax. These were themselves rubbed together or a single fibre placed between both at the appropriate angle and then withdrawn.

Owing to the difficulty of completely descaling a wool fibre without greatly reducing its strength, the experiment with descaled fibre was done with human hair. The hairs were descaled by drawing between the thumb and very fine emery paper or by attaching a small weight to one end and pulling up and down over the edge of broken glass. The fibres were regarded as sufficiently descaled when unidirectional movement no longer resulted on rubbing between a damp finger and thumb. There was no difference in the behaviour of the fibres descaled

in these two ways. The fibre was drawn between two locks lying at right angles to it and carried a negative charge whether drawn from root to tip or tip to root.

Piezo-electricity. A clean dry lock of wool, 64s Cape merino, was drawn into a half-inch internal diameter glass tube in which it was very tightly packed. One end of the tube was connected to a filter pump and the tube put in an air oven. When the whole had come to the temperature of the oven, the open end was immersed in a basin of just molten shellac. The vacuum was maintained until the wool was completely saturated with shellac, when the tube was disconnected and removed from the oven. Before it was quite cold, the glass tube was broken away and the stick of oriented wool and shellac cut with a sharp knife, at right angles to the fibre length, into pieces of convenient length.

A piece about $\frac{1}{4}''$ long, when cold, was compressed between two small brass plates. A force of about ten pounds produced a potential of the order of a volt between the plates, the root end of the fibre becoming negative on compression.

Another piece $1\frac{1}{2}''$ long was put lengthways between two brass plates, which were hot and moistened with shellac, and the whole was allowed to cool in a clamp. When cold, a shearing force of about 100 lb. gave rise to a potential of the order of 1 volt between the plates.

To meet the criticism that the effect was due to the shellac, the experiments were repeated, using top of similar wool specially prepared by repeated gilling, so that equal numbers of fibres were tip first and root first, in place of the lock. No piezo-electric potentials were found.

Pyro-electricity. A lock of clean dry wool was drawn into a thin glass tube $2'' \times {}^{5}/_{16}''$, the ends of the wool cut close, and the tube was left until all charges had leaked away. The tube was then immersed in liquid air until ebullition ceased. On holding the ends of the tube near the terminal of the electrometer, the root end was found to be negatively charged and the tip end positively. As the wool warmed up, the charges decreased, and reversed just before room temperature was reached. The reversed charges were very weak, and are the result of leakage during the time of cooling down, etc. A potential of the order of 10 volts was produced in the liquid air.

An attempt was also made to detect pyro-electricity by a modification of a method previously employed (Martin, 1931). The material to be tested is placed on a small horizontal metal plate furnished with a long glass handle at 45° to the horizontal. The plate is lowered gently into the liquid air and the handle rotated through 180°, thus bringing the plate vertical. If the material on the plate is even moderately pyro-electric, it will adhere to the plate. Wool cut to a length of about 1 mm. failed to adhere, but a mass of scales prepared from wool by a method that will be described elsewhere adhered quite strongly. Very small particles of descaled wool also failed to adhere, showing pyro-electricity to be weak or absent in the cortical cells.

A suspension of wool scales in benzene was dried by repeatedly evaporating added quantities of benzene. The scales were suspended in petrol ether (40–60°

B.P.) after the benzene had been finally evaporated off. On stirring gently they dispersed uniformly, and in the absence of agitation would settle to the bottom in the course of a few hours. The suspension was stirred up and cooled strongly in acetone and solid CO_2, or preferably a little more strongly in liquid air. A rapid flocculation and settling occurred and was complete in a few minutes. On rapid cooling to liquid-air temperatures, the petrol ether becomes so viscous that flocculation is very slow, but on re-warming it becomes extremely rapid. At any time the scales may be redispersed by vigorous agitation.

§ 3. DISCUSSION

It is well known that it is possible to obtain charges on rubbing together two bodies of the same material (Shaw, 1928 and 1930), but in the cases reported, the charges have been due to alteration of the substance during rubbing, e.g. surface strain. It does not seem that a reversal of charge with reversal of the direction of rubbing has been previously observed.

It has not yet proved possible to prepare a stable suspension of cortical cells in petrol ether, so that, though probable, it is not certain that they are devoid of pyro-electricity. It is possible that the micelles of keratin composing the cortical cells are piezo-electric and, though parallel, their sense is random, so that no bulk effect is observable. If, however, they are not piezo-electric, there must be a fundamental difference in the structure of the cuticle and cortical cells. The commonly accepted picture of the peptide chain certainly leads one to expect piezo-electricity unless the chains are arranged alternately in opposite senses, as is suggested by Astbury (1933). The matter can perhaps be decided by energy absorption in a high-frequency electric field.

REFERENCES

ASTBURY, 1933. *Trans. Faraday Soc.* **29**, 193.
MARTIN, 1931. *Min. Mag.* **22**, 519.
SHAW, 1928. *Proc. Roy. Soc.* A, **122**, 49.
SHAW, 1930. *Phil. Mag.* **9**, 577.

ERRATUM

In Vol. **53**, Part 1, p. 50, line 2, and also under "Acknowledgement", at bottom of page, *for* S. J. Smith *read* C. J. Smith.

DEMONSTRATIONS

EXCITATION OF LUMINESCENCE BY FLAME. *Demonstrated 24 May 1940 by* L. T. MINCHIN, B.Sc., M.Inst.Gas E.

SOME substances exhibit a luminescence when in contact with flames—that is to say, they glow in a manner, and to an extent, not explicable on the grounds of ordinary temperature radiation. This was first noted, in a qualitative fashion, by Balmain in 1842, but there appears to have been no reference to the phenomenon in recent scientific literature in this country, although a number of workers in Germany in the years 1913 to 1925, and another group in America between 1920 and 1933, have published papers on the subject.

In a paper presented to the Faraday Society's Symposium on Luminescence, in 1938, I drew attention to some of this work. It was evident from the discussion that there was considerable scepticism as to whether the phenomenon ever really existed. Prof. R. Tomaschek, however, mentioned the work of Donau (1913), Tiede and Buscher (1920) and others which had escaped my notice. I have since repeated some of their experiments, and there seems now no longer any reason to doubt that this form of luminescence does, in fact, exist and can quite simply be demonstrated.

The best material to use for this appears to be boron nitride, and in this demonstration it is mounted on a grooved refractory brick by being dusted on to a thin layer of kaolin paste with which the brick is coated. If this brick is then mounted above a simple bar-burner supplied with town gas, so that the sides of the flames just skim the surface (the grooves being vertical), a bluish-green light appears on the brick which is quite different from the dull red glow of incandescence which appears some moments afterwards when the material has attained red heat. The glow is most noticeable when the material has become warmed by the flame, and the light then comes and goes as the flame touches, or is moved away from, the boron nitride crystals. The instantaneous glowing under such conditions is, in fact, the most convincing demonstration that what we are witnessing is not temperature radiation, but something rather of the character of luminescence, with the flame as exciting agent.

It is worthy of note that, as Lewis (1923) and Jevons (1923) have reported, a similar luminescence occurs when suitable substances are exposed to a stream of active nitrogen from a cathode tube; and Tiede and Schleede in the same year pointed out that many substances, including boron nitride, which are luminescent when excited by flame, also show this luminescence in active nitrogen. Another link with the better known types of luminescence is afforded by the fact that traces of activators appear to be necessary. Thus, Tiede and Tomaschek (1925) showed that a trace of carbon was necessary in boron nitride, while Paneth and Winternitz

(1918) found that a trace of bismuth caused lime to luminesce blue, whereas a trace of manganese compound caused the glow to be yellow.

The probable explanation of this phenomenon is that in the flame gases there are atoms and atomic aggregates which possess an exceptional amount of energy that they are unable immediately to liberate. The conjunction of the flame with a crystal lattice of suitable structure would then enable this energy to be released in the form of light radiation and the process of combustion would be accelerated. The other suggestion—that ultra-violet radiation from the flame is the immediate exciting agent—appears less probable. Some substances which are very responsive to ultra-violet are not affected by flame, and in any case the very low intensity of this radiation in a gas flame hardly seems adequate to account for the facts.

Although it is too early to forecast possible applications to the technique of gas utilization, particularly in the lighting field, it is evident that the phenomenon of luminescence may well repay further investigation. It may be mentioned that research on this problem is now being conducted by the Arthur Duckham Research Fellow of the Institution of Gas Engineers, in collaboration with the Gas Research Board. A first report of this work, by E. C. W. Smith, has just been published (1940).

REFERENCES

BALMAIN, W. H., 1842. *Phil. Mag.* **21**, 270.
BALMAIN, W. H., 1843. *Phil. Mag.* **22**, 467.
DONAU, J., 1913. *Mh. Chem.* **34**, 949.
JEVONS, W., 1923. *Nature, Lond.*, **111**, 705.
LEWIS, E. P., 1923. *Nature, Lond.*, **111**, 529.
MINCHIN, L. T., 1939. *Trans. Faraday Soc.* **35**, 163, with addendum ; 1940. *Ibid.* **36**, 505. See also *Trans. Illum. Engng Soc., Lond.*, **5**, 75.
PANETH, F. and WINTERNITZ, E., 1918. *Phys. Ber.* **51** (2), 1728 (see p. 1736).
SMITH, E. C. W., 1940. *Inst. Gas Engrs*, no. 237.
TIEDE and BUSCHER, 1920. *Phys. Ber.* **53** (2), 2206.
TIEDE and BUSCHER, 1923. *Naturwissenschaften*, **11**, 765.
TIEDE and TOMASCHEK, 1925. *Z. anorg. Chem.* **147**, 111.

THE APPLICATION OF MICROFILM TO RECORDS OF SCIENTIFIC LITERATURE, AND A SUITABLE PROJECTION APPARATUS FOR VIEWING SUCH RECORDS. *Demonstrated* 18 *February* 1941 *by* E. LANCASTER-JONES (Science Library, Imperial College) and B. K. JOHNSON (Imperial College).

§ 1. Microfilms for scientific records, by E. LANCASTER-JONES

RESEARCH workers and students in science are sufficiently familiar with the advantages of " natural-size " photographic reproductions of scientific records in the form of photostats.

These advantages may be briefly summarized as enabling the individual to compile dossiers or files of information relating to a particular subject in a

convenient form of uniform size and with extraneous matter excluded. The original records will usually be bound up in a series of volumes containing a far greater bulk of irrelevant material. Moreover, the originals may only be accessible with great difficulty, at great expense, and for a limited time ; or may, indeed, be inaccessible.

By comparison with photostats, microfilms present the following additional advantages:—

1. Reduced bulk; giving the possibility not merely of housing within a small compass the complete file of records relating to a particular subject or project, but also facility for transporting the material and even producing it at will, by using a standard camera employing standard-gauge film.
2. Reduced cost; bringing the reproduction within the compass of the average student's purse.

As against these obvious advantages, the microfilm has only one serious drawback, namely, that it involves the employment of a special reading device of the microscope or projector type. On the other hand, it is obviously possible, and by no means impracticable, to reconstitute a natural-size positive print from a microfilm negative, at a nett cost less than that of a photostat positive; indeed, the cost is comparable with that of a photostat negative, and the product has the undoubted merit of a positive.

However, the employment of microfilm in its natural form, as a miniature of the original, presents so many advantages that the provision of facilities for reading the microfilms has become the crucial problem, upon the successful solution of which depends the transformation of the employment of microfilms in scientific study and research from a relatively rare phenomenon to a commonplace event.

There are available, though in very restricted numbers in this country, a variety of projectors adapted to the manifold needs of users of microfilms. But the cost of these projectors is prohibitive, except to institutions dealing with microfilm in large bulk.

The equivalent of the student's microscope is an obvious necessity for the individual student or research worker, and the apparatus presently to be demonstrated is designed to meet that particular need. It is adapted to provide a legible image of any section of a strip of microfilm such as is produced when the average scientific article is photographed on standard 35-mm. film, with a reduction of some fifteen diameters. Such a strip represents the normal " unit " for the purposes of scientific record, and it is easily consulted by means of the apparatus.

In conclusion, it may confidently be claimed that, whatever may be the future developments of the art and practice of producing microfilm copies of documents, an apparatus of this type will always form a convenient and essential tool in the study or laboratory of the scientific student or worker.

§ 2. A microfilm reader, by B. K. Johnson

The microfilm "reader" consists of a simple form of projection apparatus, but with modifications to meet the particular requirements of its purpose. The first of these is that there shall be sufficient light on the viewing screen to enable one to use the instrument in a room lighted either by daylight or artificial light (for it is not always convenient to have a room in complete darkness whilst reading). This involves the use of a fairly high-power source of light and suitable screening of the viewing surface. Secondly, the projection lens must give a well-defined

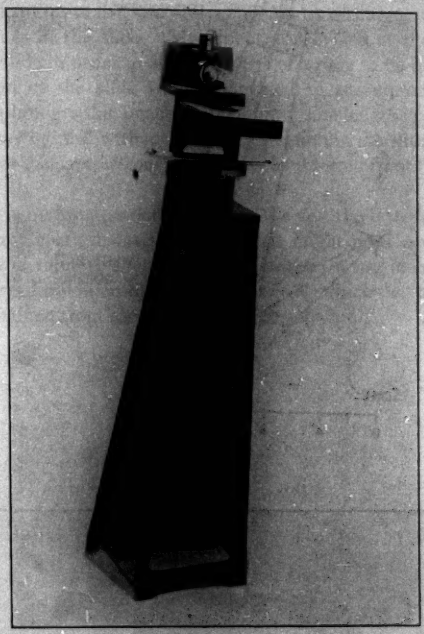


Figure 1.

image, a flat field, and freedom from distortion over an area sufficiently large to take in either one whole page of reading matter or at least one half-page. The former corresponds to approximately 20×15 mm. on the film. The magnification must also be large enough to make "comfortable" reading possible. (A suitable magnification is found to be about 16 times.)

Thirdly, the apparatus must not be costly, so that it may be available to a greater number of individuals.

In order to comply with these conditions the instrument developed is shown in figure 1 and diagrammatically in figure 2. A wooden framework, suitably inclined, provides the housing of the viewing screen at the base and serves as a

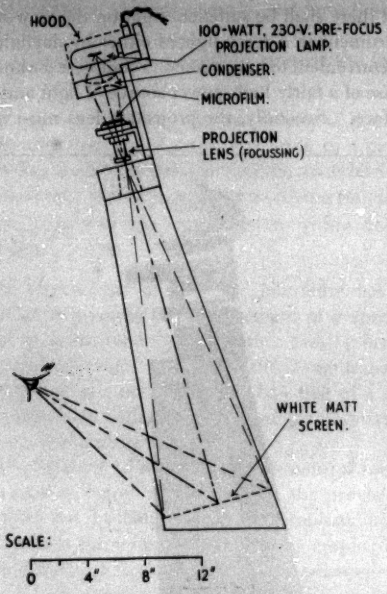


Figure 2.

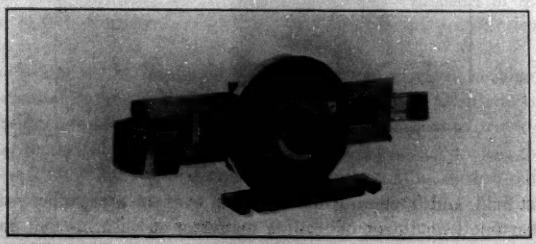


Figure 3.

support for the optical components at the top. The latter comprise a 100-watt 230-volt pre-focus projection lamp, a condenser, and the projector lens., This lens-system must be well designed and should be of 1½" to 2" focal length, working

at a fairly high F-ratio, in order to utilize as much light as possible from the source. Certain available standard lenses may be used for this, such as those employed for small cine projectors, provided a field not exceeding 19° is used, or a microscope objective as used in macrophotography. It is of interest to note that in the latter case, even a 12-volt, 36-watt standard motor-car headlamp serves quite well as a light-source, except that this requires a transformer for use with the ordinary lighting voltage, whereas the first-named lamp can be run directly from the mains.

The only other part which calls for comment is the mechanical slide for carrying the film (standard 35 mm.). This is shown in figure 3. It is essential that the film be kept quite flat, and this is ensured by placing a strip of glass over the film when it is in the slide. Movement is provided at right angles to the direction of the slide in order to allow all parts of a page to be brought on to the screen. Rotation of the slide is also provided to accommodate all possible positions of the photographic image on the film. The length of the slide is such that ten pages may be read without re-setting the film in the slide.

Dimensions of the complete apparatus may be inferred from the scale drawing of figure 2.

Whilst there is nothing particularly original about the instrument (except for simplicity and low cost), it was thought that it might meet a widely felt need and prove of some value to the individual scientist for use in his own study or laboratory for reading microphotographic copies which he may have made personally or obtained from some existing source of supply.

REVIEWS OF BOOKS

Acoustics, by ALEXANDER WOOD. (*The Student's Physics*, Volume II.) Pp. xvi + 588 (London: Blackie and Co., Ltd., 1940.) 27*s*. 6*d*. net.

The sixth of the volumes of the series entitled "The Student's Physics" maintains the exceptionally high standard set by the earlier ones. This is no re-hash of familiar text-books, but a completely new start, with the space allotted to different subjects determined entirely by their present importance, worn-out parts of the subject scrapped, and recent developments copiously drawn on for illustration. Moreover, as those who have heard Dr. Wood lecturing would expect, care has obviously been taken at every point to see that the presentation is worthy of the matter with which it deals. Perhaps the one place where re-consideration from this point of view might be desirable is on pages 371 and 372, where the analysis of progressive waves and stationary waves is compared, and Fourier's integral is used without any previous introduction. (Whilst blemishes are under notice, there is a slip on p. 310, where it is said that the thermal diffusivity of hydrogen is low, instead of high, and the paragraph in small type on p. 384, in the midst of a section on bowed strings, appears to deal with the struck string.)

The book opens with a chapter on wave motion, in which experiment is predominant, and proceeds to the analytical discussion of the same subject. In these chapters, the generality of wave theory is well brought out, and many unfamiliar analogies to light are mentioned. After a chapter on forced vibration, there is one, excellent for its modernity of outlook, on resonators, filters and horns.

Then come those familiar matters—dissipation of energy in sound waves, reflection and refraction, interference, diffraction and the measurement of velocity. Scattered through these chapters are various references to, and accounts of, modern applications to navigation, sound-ranging, detection of icebergs or submarines and other subjects. In the chapter on velocity I found out for the first time why the error in the Newtonian formula for the velocity of sound always receives so much notice. It was not so much that Newton overlooked the need for the factor γ in the formula $V=\sqrt{(\gamma p/\rho)}$, as that he sought unsuccessfully for an explanation of the non-validity of his formula, and that Euler and Lagrange both discussed it and failed to find the answer. The correction made by Laplace thus represented a bigger achievement than the student generally realizes.

The chapters on intensity and on pitch and frequency are notable for their up-to-date character, and this applies still more strongly to the following one, on analysis of sound, where quite a good introduction to the phonetics of speech is introduced.

There is not much new to say about the vibrations of strings, but the chapter on organ pipes has now to start with edge-tones and Kármán vortices. A chapter on rods, membranes and plates leads to one on the *Ear and Hearing*, which includes modern work on noise, on masking of sounds, the origin of combination tones and cognate matters.

The final chapters are on applied acoustics: recording and reproduction of sound, and acoustics of buildings. In the former, all the main recording methods are described, and a good account given of the chief types of microphone. In the latter, the Sabine formula, and the Eyring and Franklin modifications, are deduced, and their

respective fields of validity noted. There is an excellent discussion of the optimum reverberation time for rooms intended for different purposes, data on absorption coefficients of common materials, and an account of the methods by which noise transmission in buildings may be reduced.

The student who uses this book will not only have all he needs for examination purposes; he will find it interesting in itself. But its readers should not be confined to those about to sit for examinations. Many others will find it interesting, and even enjoyable, reading.

J. H. A.

Quartz Oscillators and their Applications, by P. VIGOUREUX. Pp. vi+131 (London: H.M. Stationery Office, 1939.) 4*s.* 6*d.*

Most physicists know the general facts about pressure- or piezo-electricity, but comparatively few have a precise or a detailed knowledge of the subject. In courses of instruction it is still regarded as being a very specialized part of the subject of electricity. But knowledge of this phenomenon has advanced so much in the last ten years and has so many different applications that it is time that a work appeared of the nature of a text-book suitable for both pure scientists and for those immediately interested in the technical aspect of this subject.

This book by Dr. Vigoureux, which replaces an earlier monograph entitled *Quartz Resonators and Oscillators*, is a volume which will be welcomed by both these classes of physicists.

In the chapter on the applications of the subject, some examples of which are to Wavemeter Calibration, Control of Transmitters, the Oscillator Clock, and to Ultrasonic Research, the reader gets an idea of the ramifications of this study into the general subject of physics.

The book is admirably concise and simply written. That does not mean that the book is elementary in the sense in which that term is usually applied ; it means that the author has made the subject readable.

It is a work suitable for those who have the interest of the scholar in the subject of physics, to those who have to teach the subject and require the advice and direction that a specialist can give to the research worker, and to the student.

The work also contains a list of references selected by the author from an already vast literature, from his own experience in frequent consultations of them.

H. T. F.

Atoms in Action. The World of Creative Physics, by GEORGE RUSSELL HARRISON. Pp. x+370. (London: George Allen and Unwin, Ltd., 1940.) 12*s.* 6*d.*

"Energy is wealth". This statement, quoted from the opening pages of this thrilling book, is the keynote to all that follows. "Energy is wealth", says Professor Harrison, and then proceeds to demonstrate how civilization has become enriched by the results of the research laboratory applied to industry. And always, throughout the book, the author maintains that nice balance between technological improvement and economic advantage. It may, for instance, seem cheaper at first sight to obtain electrical energy from water power than from coal, but if it so happens that the electrical energy is not required in the immediate vicinity of, say, Niagara Falls, one then has to consider which is the cheaper to transport, electrical energy or chemical energy in the form of coal. "The cheapest way to accomplish some end often is not the most economical."

To give the impression, however, that this book is concerned only with energy or with wealth would be entirely false. Professor Harrison has, in fact, given a most comprehensive survey of developments in many branches of applied physics, a survey which cannot but make one realize, with not a little awe, how much progress has been crowded into the short space of thirty or forty years. And yet throughout, like a thread through a garment, runs that persistent thought *the increased wealth of mankind*, a real wealth based on power, the power of Nature made available by scientific research and achievement. "Experience has shown no better way of eliminating poverty than by well-directed 'atom smashing'. Poverty can best be abolished by replacing it by wealth; and the systematic investigation of matter and energy without regard to immediate practical ends has turned out to be the most direct road to social riches. In the long run digging for truth has always proved not only more interesting, but more profitable than digging for gold. If urged on by the love of digging, one digs deeper than if searching for some particular nugget. Practicality is inevitably short-sighted, and is self-handicapped by the fact that it is looking so hard for some single objective that it may miss much that Nature presents to one who is purposefully digging for whatever may turn up."

One gains the impression as one turns the pages of this book that the results of many hours of painstaking investigation have been drastically condensed into a few sentences or a paragraph. Professor Harrison possesses a wide and intimate knowledge of industrial applications of physics, and a diversity of new aspects is presented to a reader who may have been more closely associated with academic than with industrial science. Not a sentence or a word in this book has been wasted, and yet reading is smooth and far from heavy, a wealth of knowledge being combined with a sparkling sense of humour and a gift for devising the aptest possible analogue. The many anecdotes too are on the same high plane. Thus we read that "in a wooden house the air at the ceiling of a room may be 25° F. warmer than at floor level. One eccentric German physics professor gave up hovering about his porcelain stove in winter; he piled tables one upon the other, climbed atop the pile, and wrote his books perched as close to the ceiling as he could get." Then, to pick another at random, there is the story of the scientist with a home recorder who produces string quartet recordings by himself. "He first plays and records the cello part. Then he plays the resulting record through on a reproducer while he accompanies its playing with another part, say that of the viola . . . ." and so on "until all the parts of the quartet have been accumulated".

Many of Professor Harrison's statements are most penetrating. Thus, "Modern life is being revolutionized by so many agencies that it seems prosaic to add another to the list; but as the steam engine introduced an industrial revolution, so the electronic vacuum tube seems to be inaugurating an equally far-reaching change in our habits of living . . . . Without such tubes we would have no x-rays, no radio broadcasting, no long-distance telephone, no talking motion-pictures, no high-fidelity recorded music." And again, "When science set the electron free it took a great step forward toward freeing mankind from its yoke of want, hunger and disease. The ransomed electron has already contributed in a thousand ways, and if the further opportunities it offers are properly utilized, will contribute in countless further ways to the comfort, safety and security of human beings." And lastly, in rather lighter vein, "A member of a majority of the Supreme Court is never wrong in legal matters, because, in the absence of higher authority, it is necessary to call him right by definition. A scientist need never be called right by definition, because he has the priceless heritage of working in a field where truth itself can be uncovered."

The titles of the chapters of this book are thrilling and original in themselves. "When Physics goes Farming", "Glass—more precious than Rubies", "Light for a living World" and "Man climbs the Winds" are but four of sixteen almost equally attractive titles.

In the course of a charming book there is only one statement (on p. 245) which jars. This statement, with regard to the attitude of the public to the development of television, is in such obvious contradiction with Professor Harrison's later views with regard to the development of the phonograph (see p. 271) and is in itself so impossible of substantiation that one may most charitably attribute it to hearsay and count it the author's one, and only, fall from grace. Moreover, in refraining from being more specific in this one criticism, the curiosity of readers of this review may be so aroused that they will be encouraged to procure a copy of the book itself, and in so doing they will be treating themselves to many hours of most delightful reading! W. B. M.

The Observer's book on Astro-Navigation, by FRANCIS CHICHESTER. Part 1, pp. xii + 103; Part 2, pp. viii + 83. (London: George Allen and Unwin, Ltd., 1941.) 2*s*. 6*d*. each part.

The subject of the navigation of aircraft by observation on celestial objects is one of increasing importance in the light of the much greater flying range of modern aeroplanes and the necessity of flying by night or during weather in which the ground is not visible. The two other navigational methods generally employed (dead-reckoning, and directional wireless) are not always reliable; owing to variation in wind velocity in the first case and due to the inadvisability of using wireless at all in certain war conditions. Consequently, astro-navigation is obviously a valuable aid to the pilot, but it must be admitted that the methods of such navigation are difficult. The author has, however, dealt with this subject in a clear and concise manner and has compressed a considerable amount of essential information into two small volumes. The limitations and accuracy of the method are discussed with frankness, and examples are worked out to illustrate the calculations involved after observations with the air sextant have been made. This is well set out and helpful to the reader. Such calculations, however, might cause even an experienced navigator a certain uneasiness as to the possibility of making errors, and the time taken to reach the final result might be considerable. On the other hand, continual practice in the routine (set out in the text) for observation and calculation may reduce the time factor, more especially if mechanical calculators are employed. The book can be recommended as a useful contribution to a subject which must of necessity occupy the minds of serious navigators at the present time, but whether the methods are too involved for the training of large numbers of men in this subject time alone will show. B. K. J.

Electron-Inertia Effects (*Cambridge Physical Tracts*), by F. B. LLEWELLYN. Pp. viii + 104. (Cambridge: The University Press, 1940.) 7*s*. 6*d*.

The general editors of *Cambridge Physical Tracts* are to be congratulated on their choice of author for the tract on electron-inertia effects. This notoriously difficult subject has been investigated by Dr. Llewellyn for some years, and his contributions to the literature have stamped him as one of the most competent workers in this field. Electron-inertia effects relate to those phenomena found in a thermionic tube on short

radio waves when the inter-electrode times of transit of the electrons become comparable with the time period of the oscillations which the tube is expected to produce, amplify or rectify.

This book of 104 pages is divided into eight chapters, the first, an introductory chapter correlating the phenomena in polarized media and ionized gases with those found in thermionic tubes. In the next three chapters the problem is formulated mathematically for the diode and the results applied in the fifth chapter to determining amongst other things the equivalent circuit of a parallel-plate diode at high frequencies. As first shown by Benham, a negative resistance condition is obtainable over certain frequency-bands and Dr. Llewellyn, after deriving this result, describes a diode designed specially to determine the existence of this negative resistance by observation of the oscillations produced thereby. Chapters 6 and 7 deal respectively with the application to triodes of the analytical results obtained for diodes, and with inertia effects on rectification at high frequencies. In the final chapter a short description in general terms is given of electronic oscillations in positive-grid triodes and magnetrons. No reference is made to resonance oscillations in magnetrons nor to the more recent velocity-modulated type of valve.

The book is interesting to read, particularly the latter portion, but its use for rapid reference is marred somewhat by the condensed style and by the fact that the author frequently refers to equations obtained in preceding portions of the book. The author assumes that the reader is acquainted with thermionic-tube symbols, as some are given in the second chapter without definition. These criticisms, however, are not serious, and this small book should form for some time the standard work on high-frequency phenomena in thermionic tubes.

J. McP.

Textbook of Light, by R. W. Stewart and J. Satterly. Seventh edition, revised by C. T. Archer. Pp. vii + 365. (London: University Tutorial Press, Ltd., 1941.) 7*s*. 6*d*.

Teachers of elementary students will welcome the appearance of the seventh edition of this standard text, which has been revised and put into a more attractive form by Mr. C. T. Archer.

A considerable portion of the book is occupied by the subject matter of previous editions suitably rearranged, but there are noteworthy additions, including modern treatments of photometry and spectra, and a chapter on interference and diffraction. The classified exercises and examination questions are now collected at the end of the book and a short but useful set of tables of optical constants is appended.

Prospective purchasers, seeking a book for class purposes, will require information concerning the sign convention adopted, and those who have not yet followed the recommendations of the Physical Society's *Report* will no doubt be relieved to learn that Mr. Archer is numbered among them.

The revision of a text-book is not such a simple process as may at first appear, and the reviser is to be congratulated on the manner in which he has welded the new and the old, and added modern material to a thorough treatment of more classical matter.

E. J. I.

Modern Physical Laboratory Practice, by J. STRONG, in collaboration with H. V. NEHER, A. E. WHITFORD, C. H. CARTWRIGHT and R. HAYWARD. Illustrated by R. HAYWARD. Pp. x + 642. (London and Glasgow: Blackie and Son Ltd., 1940.) 25*s*.

For this stimulating presentation of experimental technique its authors may most justifiably be given wholehearted and unreserved praise. As one turns the pages of this inspiring book, one feels continually impelled to seek out the nearest laboratory, there to try for oneself the experiments and processes which it so adequately describes and which are illustrated by so many beautiful and detailed drawings. Mr. Hayward's delightful drawings are, in fact, a feature of the book, not alone for their high standard of execution, but also for their conscientious regard to detail—a regard so deep that even the beads of sweat dripping from the toiling operator's brow are immortalized in the sketch illustrating the grinding of a 6-inch mirror.

Both as an inspiration and as a most valuable guide to method, this book would very profitably be read by every would-be research student before venturing to explore further the great unknown vistas of experimental physics.

The early chapters of the book deal with fundamental laboratory processes, i.e. glass-blowing, optical work, high-vacuum technique, evaporation and sputtering of metals and the uses of fused silica. Then from the general we pass to the more special uses of these processes and techniques as applied to the construction and use of electrometers and electroscopes, Geiger counters, vacuum thermopiles and optical instruments. A chapter on photoelectric cells and amplifiers is followed by chapters on photography and the production of high temperatures. The book is then concluded by three most interesting chapters describing experimental method, namely: "Notes on the Material of Research", "Notes on the Construction and Design of Instruments and Apparatus" and "Molding and Casting".

It is perhaps a little surprising, however, to find, in a book of such high quality, some misprints which must have been overlooked in proof-reading. Thus Van de Graaff's name is misspelt twice, A. F. C. Pollard's initials suffer the same permutation three times, the dimensions of $(1/l)\,(\Delta l/t_1-t_2)$ are obviously wrong, and one really wonders whether "spashed" bears the same relation to "splashed" as "sputtered" does to "spluttered"! But these are minor details which will doubtless be eradicated from the next edition. This, at the rate the book should sell, will not mean waiting very long! W. B. M.

RECENT REPORTS AND CATALOGUES

A Catalogue of English and Foreign Books and Periodicals on Pure and Applied Mathematics, Physics, etc., mainly works from the libraries of the late Sir J. J. Thomson, O.M., F.R.S., and the late Professor S. L. Loney, M.A. (Catalogue no. 495.) Pp. 34. BOWES & BOWES, LTD., 1 & 2 Trinity Street, Cambridge.

Abstracts of Patents relating to Rubber Latex, by T. R. DAWSON and R. W. PARRIS. (A supplement to *Rubber Latex*.) Pp. 136. Issued by THE BRITISH RUBBER PUBLICITY ASSOCIATION, 19 Fenchurch Street, London, E.C. 3. (February 1941.)

Cambridge Books, Spring 1941. Pp. 28. CAMBRIDGE UNIVERSITY PRESS, 200 Euston Road, London, N.W.1.

Foster (Resilia) Optical, Radiation and Photoelectric Pyrometers. (Book no. 73.) Pp. 54. FOSTER INSTRUMENT CO., LTD., Letchworth, Herts.

Foster (Resilia) Practical Pyrometers, Thermo-Couple (Stem) Type. (Book no. 76.) Pp. 76. FOSTER INSTRUMENT CO., LTD., Letchworth, Herts.

Roof Spotter System. (List no. 91.) LONDEX LTD., 207 Anerley Road, London, S.E.20.

The Wild-Barfield Heat-Treatment Journal (Economy Issue), vol. IV, no. 27. Pp. 23–34. WILD-BARFIELD ELECTRIC FURNACES, LTD., Watford, Herts.